243
1.9.70

CHILD DEVELOPMENT

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

REGINALD MACGREGOR ARCHIBALD
Rockefeller Institute for Medical Research

JOSEPH H. BRITTON
Pennsylvania State University

SIDNEY W. BIJOU
University of Washington

VAUGHN J. CRANDALL
Fels Research Institute

ERCEL S. EPPRIGHT
Iowa State College

C. GLEN KING
Nutrition Foundation

BENJAMIN PASAMANICK
Ohio State University

ANSELM STRAUSS
University of Chicago

RALPH L. WITHERSPOON
Florida State University

Managing Editor

KATE HOFFMAN

BUREAU OF EDNL. & PSYL. RESEARCH.
LIBRARY

VOLUME 27

MARCH, 1956 *to* DECEMBER, 1956

CHILD DEVELOPMENT PUBLICATIONS

SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.
Purdue University, Lafayette, Indiana
1956

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated 1-9-70
Acc. No. 243

CONTENTS

VOLUME 27, 1956

ANDERSON, JOHN E. Child Development: An Historical Perspective 181

BAYLEY, NANCY. Individual Patterns of Development 45

CALVIN, ALLEN D., PERKINS, M. J., & HOFFMANN, FREDRIC K. The Effect of Nondifferential Reward and Non-Reward on Discriminative Learning in Children 439

CARROLL, JOHN B., & LEVIN, HARRY. A Method for Determining the Polarity of Behavior Items 427

CASTANEDA, ALFRED, MCCANDLESS, BOYD R., & PALERMO, DAVID S. The Children's Form of the Manifest Anxiety Scale 317

CASTANEDA, ALFRED, PALERMO, DAVID S., & MCCANDLESS, BOYD R. Complex Learning and Performance as a Function of Anxiety in Children 327

CASTANEDA, ALFRED (*see also* McCandless, Boyd R.) 379

CASTANEDA, ALFRED (*see also* McCandless, Boyd R.) 385

CASTANEDA, ALFRED (*see also* Palermo, David S.) 333

CRANDALL, VAUGHN J. Some Problems of Personality Development Research 197

DEMISCH, ARTHUR, & WARTMANN, PETER. Calcification of the Mandibular Third Molar and Its Relation to Skeletal and Chronological Age in Children 459

GAIER, EUGENE L. (*see* Morgan, Patricia K.) 447

GEWIRTZ, JACOB L. A Factor Analysis of Some Attention-Seeking Behaviors of Young Children 17

GEWIRTZ, JACOB L. A Program of Research on the Dimensions and Antecedents of Emotional Dependence 205

GOLLIN, EUGENE S. Some Research Problems for Developmental Psychology 223

HOFFMANN, FREDRIC K. (*see* Calvin, Allen D.) 439

HOLDEN, RAYMOND H. The *Children's Apperception Test* with Cerebral Palsied and Normal Children 3

JACKSON, PHILIP W. Verbal Solutions to Parent-Child Problems 339

JONES, HAROLD E. The Replacement Problem in Child Development 237

KOCH, HELEN L. Children's Work Attitudes and Sibling Characteristics 289

KOCH, HELEN L. Some Emotional Attitudes of the Young Child in Relation to Characteristics of His Sibling 393

LEVIN, HARRY, & SEARS, ROBERT R. Identification with Parents as a Determinant of Doll Play Aggression 135

LEVIN, HARRY (*see also* Carroll, John B.) 427

LEVITT, EUGENE E. The Development of Prejudice: A Critique of the *Horowitz Faces Test* 155

MCCANDLESS, BOYD R., & CASTANEDA, ALFRED. Anxiety in Children, School Achievement, and Intelligence 379

MCCANDLESS, BOYD R., CASTANEDA, ALFRED, & PALERMO, DAVID S. Anxiety in Children and Social Status 385

MCCANDLESS, BOYD R., & SPIKER, CHARLES C. Experimental Research in Child Psychology 75

MCCANDLESS, BOYD R. (*see also* Castaneda, Alfred) 317

MCCANDLESS, BOYD R. (*see also* Castaneda, Alfred) 327

MCCANDLESS, BOYD R. (*see also* Palermo, David S.) 333

MARTIN, WILLIAM E. Editorial Note 83
MORGAN, PATRICIA K., & GAIER, EUGENE L. The Direction of Aggression in the Mother-Child Punishment Situation 447
MOUSTAKAS, CLARK E., SIGEL, IRVING E., & SCHALOCK, HENRY D. An Objective Method for the Measurement and Analysis of Child-Adult Interaction 109
PALERMO, DAVID S., CASTANEDA, ALFRED, & MCCANDLESS, BOYD R. The Relationship of Anxiety in Children to Performance in a Complex Learning Task 333
PALERMO, DAVID S. (*see also* Castaneda, Alfred) 317
PALERMO, DAVID S. (*see also* Castaneda, Alfred) 327
PALERMO, DAVID S. (*see also* McCandless, Boyd R.) 385
PERKINS, M. J. (*see* Calvin, Allen D.) 439
ROSE, ARNOLD M. Reference Groups of Rural High School Youth 351
ROSENBLUM, SIDNEY. The Effects of Differential Reinforcement and Motivation on Prediction Responses of Children 99
SCHAEFFER, MAURICE (*see* Shepard, Winifred O.) 173
SCHALOCK, HENRY D. (*see* Moustakas, Clark E.) 109
SEARS, ROBERT R. (*see* Levin, Harry) 135
SHEPARD, WINIFRED O. The Effect of Verbal Training on Initial Generalization Tendencies 311
SHEPARD, WINIFRED O., & SCHAEFFER, MAURICE. The Effect of Concept Knowledge on Discrimination Learning 173
SIEGEL, ALBERTA ENGVALL. Film-Mediated Fantasy Aggression and Strength of Aggressive Drive 365
SIGEL, IRVING E. The Need for Conceptualization in Research on Child Development 241
SIGEL, IRVING E. (*see also* Moustakas, Clark E.) 109
SPIKER, CHARLES C. The Effects of Number of Reinforcements on the Strength of a Generalized Instrumental Response 37
SPIKER, CHARLES C. Experiments with Children on the Hypotheses of Acquired Distinctiveness and Equivalence of Cues 253
SPIKER, CHARLES C. The Stimulus Generalization Gradient as a Function of the Intensity of Stimulus Lights 85
SPIKER, CHARLES C. (*see also* McCandless, Boyd R.) 75
WARTMANN, PETER (*see* Demisch, Arthur) 459
WENAR, CHARLES. The Effects of a Motor Handicap on Personality: III. The Effects on Certain Fantasies and Adjustive Techniques 9
WRIGHT, HERBERT F. Psychological Development in Midwest 265

CHILD DEVELOPMENT

Vol. 27
March, 1956
No. 1

Editor
WILLIAM E. MARTIN
Purdue University

Consulting Editors

REGINALD MACGREGOR ARCHIBALD
Rockefeller Institute for Medical Research

JOSEPH H. BRITTON
Pennsylvania State University

SIDNEY W. BIJOU
University of Washington

VAUGHN J. CRANDALL
Fels Research Institute

ERCEL S. EPPRIGHT
Iowa State College

C. GLEN KING
Nutrition Foundation

BENJAMIN PASAMANICK
Ohio State University

ANSELM STRAUSS
University of Chicago

RALPH L. WITHERSPOON
Florida State University

Managing Editor
KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Pudue University
Lafayette, Indiana

Made in United States of America

Ednl. Research
Dated
Accs. No

Printed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

THE *CHILDREN'S APPERCEPTION TEST* WITH CEREBRAL PALSIED AND NORMAL CHILDREN [1]

RAYMOND H. HOLDEN

At present there is fairly general agreement about some cognitive aspects of cerebral palsy (5, 6, 11). But a survey of the literature reveals little evidence of exploration of conative aspects of personality (3, 8). Besides a few reports of paper-and-pencil personality inventories (10, 12), there are only two notable studies, Klapper and Werner's (9) cerebral palsied twin study characterized by excellent methodology, and Werner's Rorschach study of brain-injured children (13) which emphasizes qualitative differences between the experimental and non-brain-injured control group.

In an attempt at an initial exploration of the personality of school-aged cerebral palsied children, the Children's Apperception Test (hereafter referred to as CAT) was selected as a projective instrument. The purpose of this study is to evaluate the effectiveness of the CAT as a projective test with cerebral palsied children.

The CAT was devised by Bellak (2) for children aged 3 to 10 years of age. The test consists of 10 pictures depicting animals in various situations. Table 1, Column 2, briefly describes the content of each picture card. The CAT is a descendant of Murray's Thematic Apperception Test, which Bellak (2) feels is unsuitable for young children. The basic assumption of the CAT is that it is an apperceptive method: "a method of investigating personality by studying the dynamic meaningfulness of the individual defferences in perception of standard stimuli" (2, p. 1). The CAT is said to be "able to reveal the dynamics of interpersonal relationships, of drive constellations and the nature of the defenses against them" (2, p. 2). It is hoped that the results of this study will shed some light on the generality of these statements as they apply to brain-injured (cerebral palsied) children.

PROCEDURE

Subjects. Eight cerebral palsied children (five boys, three girls) attending the Meeting Street School Cerebral Palsy Center, Providence, Rhode Island, composed the experimental group. The chronological age of these children ranged from 5 years, 7 months, to 12 years, 11 months, with an average of 9 years, 10 months. Their average mental age was 7 years, no months, and their average IQ was 73.1, only slightly lower than previously

---

[1] This paper was presented at the first Scientific Session of the Rhode Island Psychological Association, Providence, October, 1954.

CHILD DEVELOPMENT, Vol. 27, No. 1 (March, 1956)

reported mean IQs for large cerebral palsied populations (8). Seven of the eight children had IQs between 60 and 80, and one child's IQ was 116.

Seven normal school children and three children with behavior problems constituted the control group. The results were pooled because inspection of the data revealed no noticeable differences in the variables under consideration. The finding of no observable differences between stories given by normal and neurotic adult subjects on the TAT, for example, has already been reported in the literature (1). The seven "normals" were five boys and two girls from the Walter J. Paton Elementary School in Shrewsbury, Massachusetts, and the three "neurotic" children were two boys and a girl from the Emma Pendleton Bradley Home, Riverside, Rhode Island, an inpatient hospital for disturbed children. The average chronological age of the control group was 7 years, 3 months (range: 6 years, 2 months, to 7 years, 8 months). The average mental age was 7 years, 5 months (range: 5 years, 10 months, to 8 years, 2 months), and their average IQ was 101.5 (range: 92 to 115).

It can be seen that the experimental and control groups are most nearly equated for mental age. Thus the cerebral palsied children, although having had longer experience in the world, are nearly equal to the control group in level of mental ability, although not equal in the rate of mental growth (as evidenced by the widely discrepant average IQs).

The CAT was administered individually by the author according to standard administration procedure described by Bellak (2). It was the author's qualitative impression that motivation and interest of subjects in

TABLE I

ANALYSIS OF CAT STORIES BY CARDS

		CEREBRAL PALSY		CONTROL	
Card	*Most Common Description*	*Description*	*Thema*	*Description*	*Thema*
I	Chickens eating	100%	0%	30%	70%
II	Bears pulling on rope	75%	25%	10%	90%
III	Lion sitting on chair	50%	50%	30%	70%
IV	Kangaroo (or other animal) on bicycle	50%	50%	20%	80%
V	Bears(s) sleeping	50%	50%	50%	50%
VI	Bears sleeping	87%	13%	30%	70%
VII	Lion (tiger) chasing monkey	62%	38%	40%	60%
VIII	Four monkeys (gorillas) eating (talking)	62%	38%	20%	80%
IX	Rabbit in bed	38%	62%	20%	80%
X	Two dogs in bathroom (spanking)	38%	62%	30%	70%
	Average Per Cent Descriptions	61%		28%	
	Average Per Cent Thema		39%		72%

both groups were high. All children responded with apparent pleasure to the pictures, and no card was ever refused by any child.

Analysis of data. Test data were analyzed in two ways: (1) the percentage of card description without story content (Table 1); and (2) the number of descriptions per subject in both the cerebral palsied and control groups (Table 2).

TABLE 2

ANALYSIS OF CAT DESCRIPTIONS BY SUBJECTS

GROUP	NUMBER OF DESCRIPTIONS										
Subject:	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Average*
Cerebral Palsy	7	9	5	5	1	8	7	6	–	–	6.0
Control	2	0	8	6	2	0	0	2	8	0	2.8

The classification of CAT stories as description or thema was a fairly easy one. All stories were considered as "description" that contained only the elements found in the most common description of each card with no story added. A "thema" is defined as any description of the card, *plus* the verbalization of additional story material, feeling, tone, affect, or ending to the story which could be interpreted according to some personality theory.

A test of reliability of classifying the stories was done. A random sample of 50 per cent of the records was classified independently by another psychologist. Eighty-five per cent agreement was obtained on cerebral palsied stories, and 93 per cent agreement on the classification of the control group stories.

To give a flavor of the difference between descriptions of the cerebral palsy group and the thema of the control group, the following examples are cited as illustrative. First, a thema given by a normal subject:

"Once there was a mother kangaroo and a baby kangaroo. Baby kangaroo and mother was going away to the circus. The baby couldn't ride his bicycle any more. He had to walk. Baby kangaroo was hitting his mother with a balloon so they took the balloon away from baby and the balloon accidentally broke and baby ran after the balloon. The baby climbed a tree and caught the balloon and he just got it and it hit a sharp rock and it broke and they went to the circus and he asked the man for a balloon and he got it" (Subject N2, Story 4).

Length is not important, but expression of affect is, as witnessed by the short thema on Card 6: "Seems to be a mommy and poppy bear and a little baby bear hiding in a cave. Looks like they're frightened." (Subject N5, Story 6).

By comparison now, here are descriptions, first by a normal: "Three chickens are going to eat all the porridge." (Tell a story?) "They'll eat all the porridge." (Subject N5, Story 1). Here is nothing but the bare essentials, only a description commonly given. Second, a story classified "description" by a brain-injured subject: "The baby goat's riding a bike and the mother goat is carrying a basket with food and baby's carrying a balloon." (Anything else in the story?) "No." (Subject CP2, Story 4).

Something which the non-brain-injured never did, but which was done by 15 per cent of the brain-injured group, was to get so involved in the actual details of the picture that mere enumeration of small details took place. This is the phenomenon of concrete stimulus - boundedness described by Goldstein (7) and Strauss and Werner (11). The following is an example:

"This could be the monkey and this could be the tiger running up to the monkey to eat him. This could be the tree, this could be flowers, this could be the bottom, and this could be the back, this could be the leg, this could be the tail, this could be the water." (Subject CP8, Story 7).

Use of irrelevant details, ordinarily overlooked by the majority of subjects, but made to fit into the story occurs only in brain-injured subjects:

"This could be the father dog and this could be baby dog. In the bathroom. This could be the hanger for the towel, this could be the toilet, this could be the top of the toilet, this could be the chain." (What is happening?) "Baby dog is eating." (Why?) Subject points to the dog with mouth open. "See, he's hungry." (Subject CP8, Story 10).

RESULTS

There is a striking difference between the number of descriptions given by cerebral palsied and non-brain-injured children. Table 1 shows, for example, that cerebral palsied children gave descriptions only to 61 per cent of the cards. Normals, on the other hand, gave only 28 per cent descriptions, and 72 per cent thema. This is a significant difference between groups beyond the .01 level of confidence.

Table 2 analyzes the results by individual subjects. The cerebral palsy and normal subjects are not paired, but numbered in random order. Considering the results from this table, it can be seen that the cerebral palsied child gives on the average 6.0 descriptions out of a possible 10. The normal child, on the other hand, gives less than half this number, on the average 2.8 descriptions out of a possible 10.

However, if we look at the individual cases more closely, we find that in the control group three cases contribute greatly to distorting the average number of descriptions given. These cases are Case 3 (eight descriptions), Case 4 (six descriptions), and Case 9 (eight descriptions). These cases were further investigated by the author. Cases 3 and 4 were children in public school with average intelligence and without psychological problems for over one year (as determined by the initial requirements for inclusion in this control group). Nevertheless, when their teacher was questioned more closely, she admitted that although the children were no problem, they were both "strange" and sometimes had difficulty in coordination. It was not possible, despite further efforts, to obtain the necessary medical records or history to ascertain any previous brain injury. Case 9, a boy at the Emma Pendleton Bradley Home, was subsequently found to have a diagnosis of cerebral palsy, mild, secondary to behavior disorder.

This evidence, if it can be substantiated by validation study, appears to mean that the CAT, in terms of whether predominantly descriptions or predominantly thema are given, can distinguish between brain-injured and non-brain-injured children in the particular age range and mental age range studied. Intelligence does not appear to be a significant factor. Cerebral palsy Case 1, a boy with an IQ of 116, brighter than any subject of the control group, nevertheless, gave 7 out of 10 descriptions to the cards in the typically brain-injured manner.

Theoretical Implications

Theoretically, these qualitative differences between groups provide evidence for a discontinuity theory of genetic psychology. The brain-injured child's performance is qualitatively different from the non-brain-injured, and is on a lower more primitive level despite similarities of mental age.

Cerebral pasied children, like the organic adults described by Goldstein (7), show more concrete behavior than abstract, and more concrete behavior than non-brain-injured children of their same mental age. They appear relatively unable to adopt an abstract attitude to "tell a story," are stimulus-bound, and rigidly continue with the concrete attitude giving descriptions to the cards. This study corroborates Werner's interpretations of Rorschachs of brain-injured children (13), in which he found restricted fantasy life, forced responsiveness to stimuli, meticulousness, and difficulty in shifting.

Summary

In summary, the difference in approach to the CAT between cerebral palsied and non-brain-injured children is significant in terms of phenotypically giving more card descriptions than thema and genotypically in persisting in a concrete rather than an abstract attitude. This genotypic finding is related to similar conclusions, despite different cognitive or conative tasks, by Werner (13), Strauss and Werner (11), and Dolphin and Cruickshank (5). This study gives weight to a discontinuity theory of genetic psychology.

On the basis of the present results, if confirmed by validation studies, the CAT might serve as one psychological differentiator between brain-injured and non-brain-injured children. Further research is required to determine the limits of mental ages which might be used as cut-off points to separate the two groups, the brain-injured and the non-brain-injured.

References

1. Bellak, L. The concept of projective psychology. In L. E. Abt & L. Bellak (Eds.), *Projective psychology.* New York: Knopf, 1950. Pp. 7-32.
2. Bellak, L., & Bellak, S. S. *The Children's Apperception Test Manual.* New York: C.P.S., 1949.
3. Block, W. E. Personality of the brain-injured child. *Except. Child.,* 1954, 21, 91-100.
4. Cruickshank, W. M., & Dolphin, J. E. The emotional needs of crippled and non-crippled children. *Except. Child.,* 1949, 16, 33-40.
5. Dolphin, J. E., & Cruickshank, W. M. The figure-background relationship in children with cerebral palsy. *J. clin. Psychol.,* 1951, 7, 228-231.
6. Dolphin, J. E., & Cruickshank, W. M. Visuo-motor perception in children with cerebral palsy. *Quart. J. child Behavior,* 1951, 3, 198-209.
7. Goldstein, K. *The organism.* New York: American, 1929.
8. Holden, R. H. A review of psychological studies in cerebral palsy: 1947-1952. *Amer. J. ment. Def.,* 1952, 57, 92-99.
9. Klapper, Z. S., & Werner, H. Developmental deviations in brain injured (cerebral palsied) members pairs of identical twins. *Quart. J. child Behavior,* 1950, 2, 288-313.
10. Portenier, L. G. Psychological factors in testing and training the cerebral palsied. *Physiotherapy Rev.,* 1942, 22, 301-303.
11. Strauss, A. A., & Werner, H. Disorders of conceptual thinking in brain injured child. *J. nerv. ment. Dis.,* 1942, 96, 153-172.
12. Tracht, V. S. A comparative study of personality factors among cerebral palsied and non-handicapped persons. Unpublished master's thesis, Univer. Chicago, 1946.
13. Werner, H. Perceptual behavior of brain injured, mentally defective children: an experimental study by means of the Rorschach technique. *Genet. Psychol. Monogr.,* 1945, 31, 51-110.
14. Werner, H. *Comparative psychology of mental development.* Chicago: Follett, 1948.

THE EFFECTS OF A MOTOR HANDICAP ON PERSONALITY: III. THE EFFECTS ON CERTAIN FANTASIES AND ADJUSTIVE TECHNIQUES[1]

CHARLES WENAR
University of Illinois, College of Medicine

Investigators of the personality characteristics of children with motor handicaps have emphasized the many emotional problems which arise in connection with such physical disabilities. Frequently mentioned are the effects on the child's body image, the special complications of dependency needs and the expression of hostile impulses, parental reactions of guilt, rejection or overprotection, and the peculiar social attitudes of the non-handicapped toward physical disability. The diverse and sometimes contradictory experimental findings also indicate the exceedingly complex nature of such problems. The present series of studies was designed to explore certain aspects of this general area.

STATEMENT OF THE PROBLEM

The following aspects of the handicapped child's emotional life were investigated: (1) his over-all adjustment; (2) the variety of his interests; (3) his fantasies of danger and destruction; and (4) his self-protective defenses. Since a different methodological approach was used, each aspect will be presented individually.

METHOD AND RESULTS

Subjects and Experimental Task

Data from the World Test were obtained from a group of 12 non-handicapped, 12 moderately handicapped, and 12 severely handicapped children. All were pupils in the Chicago public school system, between the ages of 8 and 10, of average or above intelligence, with no other severe physical handicap. Further details concerning the sample can be found in

---

[1] These studies were sponsored by The Coordinated Program for Handicapped Children and were done at The Institute for Psychosomatic and Psychiatric Research and Training, Michael Reese Hospital, Chicago, Ill.

CHILD DEVELOPMENT, Vol. 27, No. 1 (March, 1956)

the first paper of this series (5), while a description of the administration of the World Test to these subjects is given in the second paper (6).

Over-all Adjustment

One can hypothesize that the increased number of problems encountered by a child with a motor handicap will increase the probability of his becoming maladjusted. Thus it is possible that, as a group, the handicapped children will have a higher incidence of emotional disturbances of one sort or another than the non-handicapped children.

In order to test this hypothesis a measure of over-all adjustment was obtained by using the signs of emotional disturbance developed and validated by Buhler, Lumry and Carrol (2). These signs are: Aggressive Worlds (all play with aggressive content such as accidents, killing), Empty Worlds (less than 50 pieces,[2] less than 5 categories, no men and women used), Closed Worlds (the entire World or large parts fenced in or enclosed), Rigid Worlds (marked by exaggerated symmetry), and Disorganized Worlds (pieces placed chaotically). Each World was scored for the number of signs it contained and the mean number of signs for each group was calculated. These means are as follows: Non-Handicapped (N-H) Group—1.33 signs; Moderately Handicapped (M-H) Group—1.50 signs; Severely Handicapped (S-H) Group—1.17 signs. An analysis of variance yielded an *F* of .27 which fails to approach significance at the 5 per cent level of confidence.

However, Buhler *et al.* also claim that "the presence of two or more of these symptoms is highly discriminatory of abnormality" (p. 34). Therefore the number of children in each group with less than 2 signs and the number with 2 or more signs was calculated. This proportion was 7 to 5 in the N-H Group, 7 to 5 in the M-H Group, and 8 to 4 in the S-H Group. Since the proportions were so similar no statistical test of significance was deemed necessary.[3]

Therefore, there is no evidence that a motor handicap, per se, leads to a significantly higher incidence of general maladjustment than is found in non-handicapped children.

Variety of Interests

One can hypothesize that a motor handicap will produce a constriction of interest and a retreat from active involvement in interpersonal relations. This may be the consequence either of a realistic narrowing of the handicapped child's range of experiences or of a protective withdrawal into simple, easily mastered activities.[4]

---

[2] This sign was omitted because the physical limitations of the handicapped groups forced them to use significantly fewer pieces.

[3] In Buhler's study the Normal group was selected on the criterion of good adjustment, whereas in the present study the Non-Handicapped children were selected without regard to their adjustment. Therefore it is not surprising to find a certain per cent of them falling into the maladjusted category.

Decrease in the variety of objects used in constructing the World was used as one index of a narrowing of interests. To obtain a measure of this, the number of categories of pieces was computed for each group. Bolgar's categories (1) were used and include Houses, Public Buildings, Trees, Fences, Males, Females, Special People, Soldiers, Domestic Animals, Dogs, Wild Animals, Bridges, Vehicles, Details, Boats and Planes. The mean number of categories for each group was calculated and were as follows: N-H—13.08; M-H—13.33; S-H—10.67. An analysis of variance yielded an *F* of 1.75 which fails to approach significance at the 5 per cent level of confidence.

Therefore there is no evidence of a decrease in the variety of interests in the handicapped groups. This finding becomes more impressive when one remembers that there is a significant decrease in the total number of pieces used as the motor handicap becomes increasingly severe.

TABLE I

INTRODUCTION OF HUMAN PIECES

Group	*First Section*	*Second Section*	*None Used*
N-H	8	2	1
M-H	7	2	1
S-H	8	2	2

One might expect a retreat from active involvement in interpersonal relations to be expressed by a shying away from the use of people in constructing the Worlds. Thus the percentage of human pieces was computed and group means obtained. These means were as follows: N-H—16.33; M-H—21.83; S-H—21.83. An analysis of variance yielded an *F* of .65 which fails to approach significance at the 5 per cent level of confidence.

As a more refined approach to this problem, the data were examined in terms of the point at which human pieces were introduced into the construction. Thus the children with difficulties in interpersonal relations should introduce such pieces later in their construction than the less inhibited children. In order to obtain a measure of this factor, construction of the World was divided into thirds and the test was scored as to whether humans were introduced in the first, second, or third section. Since no child introduced such pieces in the third section, but some did not use human pieces at all, the category "None used" was substituted for the third section. The number of children in each group who introduced human pieces at the different stages of construction are presented in Table 1.[5] Because the proportions were so nearly equal, a chi square was not deemed necessary. Thus there is no evidence that children with a motor handicap retreat from interest in interpersonal relations.

---

[4] The author is grateful to Dr. Hedda Bolgar for suggestions as to the analysis and interpretation of the World Test in the remaining sections of this paper.

Fantasies of Destruction

One can hypothesize that a motor handicap would increase a child's concern over destructiveness. It can be reasoned that the frustrations he experiences could intensify his own hostile, destructive impulses, and that his helplessness in many situations would lead to an exaggeration of the destructive forces in his environment.

To test this hypothesis, the Worlds were examined for all themes which showed concern over destruction. Such themes included interpersonal violence (war, fighting, arguments), fires, accidents, sickness, zoo or jungle scenes involving fighting, biting, or escape of wild animals. To be scored, the theme either had to be verbalized during the enquiry as a destructive theme per se or as a singling out of pieces with dangerous or destructive implications, or it had to be clearly indicated by the structure, such as elaborating the hospital with trees and a doctor going in the door. The pieces most frequently used in such themes were soldiers, wild animals, fire house and engines, hospital, doctor, sailor and policeman, although individual children showed great ingenuity in devising novel situations (e.g., a tree doctor coming to repair a diseased tree). The percentage of such pieces in the World was used as a measure of the child's concern with themes of destruction. The mean per cents of these destructive pieces were as follows: N-H—7.41; M-H—12.16; S-H—6.33. An analysis of variance yielded an F of 2.25 which fails to approach significance at the 5 per cent level of confidence.

However, inspection of the data indicated that in the handicapped groups themes of destruction were introduced much earlier in the construction; in fact, soldiers, wild animals and policemen were used as first pieces significantly more frequently in these groups. In order to discover whether the frequency of destructive themes differed significantly for the three groups at different stages of construction, a sequential analysis was undertaken. In this analysis the construction of the Worlds was divided into fifths, and the number of destructive pieces in each fifth was tabulated. The total number of destructive pieces in each segment is presented in Table 2.[6]

---

[5] The N in the groups does not always equal 12 because some cases were lost due to the author's failure to develop an adequate recording system for the sequential construction of the Worlds.

[6] Because of the author's failure to develop an adequate system of recording the sequential construction of the Worlds, the N in the groups are as follows: N-H, 8; M-H, 7; S-H, 10.

TABLE 2

NUMBER OF DESTRUCTIVE PIECES IN EACH SECTION

Group	*First*	*Second*	*Third*	*Fourth*	*Fifth*
N-H	3	6	5	7	11
M-H	9	3	2	2	3
S-H	2	2	6	2	2

A trend analysis yielded an F of 2.03 which approaches significance at the 5 per cent level of confidence. Thus there is a tendency for destructive themes to be used differently at consecutive stages of construction; in the Non-Handicapped group there is a gradual increase in such themes, in the Moderately Handicapped group there is an initial emphasis which tapers off quickly, and in the Severely Handicapped group the emphasis comes midway in the construction.

While there is no difference in the number of fantasies of destruction, the handicapped and non-handicapped children do differ significantly in their pattern of preoccupation with such fantasies.

Protective Defenses

One can hypothesize that the handicapped child will feel himself more vulnerable to internal and external threats and consequently will adopt rigid self-protective defenses. Thus he can isolate himself from the unpleasant aspects of his environment and protect himself from expressing dangerous impulses.

Buhler regards the use of fences in constructing a World as a measure of self-control, too many enclosures indicating high anxiety and rigid control, too few enclosures indicating impulsivity or a tendency toward disintegration. Therefore the percentage of fences was used as a measure of protective defenses. The mean per cents for each group were calculated and were as follows: N-H—30.38; M-H—20.67; S-H—15.92. An analysis of variance yielded an F of 5.49 which approaches significance at the 1 per cent level of confidence.

Therefore there is a significant decrease in the use of fences in children with a motor handicap, indicating a weakened ability to erect protective defenses against internal or external threats.

DISCUSSION

The present findings, based on the data from the World Test, indicate that a motor handicap, per se, does not increase the incidence of emotional disturbance, as there is the same proportion of adjusted and maladjusted children in the handicapped and non-handicapped groups. There is also no evidence of a significant impoverishment of interests or retreat from interpersonal relations. On the contrary, the handicapped children's interests are as varied and their concern over interpersonal interactions is as great as the non-handicapped children's. Such findings, together with those of the previous study (6) emphasize the restitutive and compensatory mechanisms which are available to the handicapped children; they suggest a successful coping with many problems and a maintenance of an active and lively interest rather than a retreat into rigidity or apathy.

However, there is evidence that the motor handicap does entail significant personality changes. Most striking is a weakening of the ability to erect defenses which would result in a more stable kind of self-control. The non-handicapped child's characteristic pattern is one of initial inhibition of anxieties and angers which only gradually come to the surface. Corroborative evidence for this comes from other studies (3, 4) in which play material is used over a number of sessions. On the other hand, the handicapped group seems to be more immediately and more chronically vulnerable to feelings of threat, either from his own destructive impulses or from a world which he perceives as dangerous. Thus one can infer that the handicapped child does not have the type of self-control which would enable him to hold negative feelings in check and function on a more objective, realistic basis; on the contrary, his concern with internal or external dangers is often immediate and persistent. This weakening of emotional objectivity parallels the weakening of objectivity in setting realistic goals found in the first study of this series (5). Both findings suggest that emotional instability rather than rigidity is a primary effect of the motor handicap on personality.

Two methodological points should be made. The above conclusions were based on data from the World test which has been validated as to its diagnostic usefulness but not as to the criteria of fantasies of danger and self-protective defenses. Therefore the results should be accepted with the qualifications reserved for such procedures. One should also be cautioned against overgeneralizing the findings because of the limited nature of the sample, both in terms of size and age range. Second, the author was impressed in this study, as in a previous one (5), by the fruitfulness of a trend analysis. Future research, therefore, might well be concerned not only with discovering group differences in a given characteristic but with the temporal patterning of such characteristics. For example, in studying dependency it might be that the non-handicapped child is initially independent and will allow himself to become dependent only after a certain period of time, while the handicapped child might expect dependent gratifications much sooner. It is possible that many personality traits in handicapped children are no different quantitatively or qualitatively, but the organization of needs and defenses is changed.

SUMMARY

Analysis of the World Test administered to handicapped and non-handicapped children revealed the following results: There was no difference between the groups in terms of over-all adjustment, variety of interests and concern over interpersonal relations; there was also no difference in the number of destructive fantasies, but the handicapped children were significantly weaker in their ability to establish self-protective defenses and were more quickly or more chronically concerned with their destructive fantasies.

REFERENCES

1. BOLGAR, H., & FISCHER, L. K. Personality projection in the World Test. *Amer. J. Orthopsychiat.*, 1947, 17, 117-128.
2. BUHLER, C., LUMRY, G. K., & CARROL, H. S. World-Test standardization studies. *J. Child Psychiat.*, 1951, 2, 1-81.
3. PHILLIPS, RUTH. Doll play as a function of the realism of the material and the length of the experimental sessions. *Child Develpm.*, 1944, 15, 123-144.
4. PINTLER, M. H. Doll play as a function of experimenter-child interaction and initial organization of materials. *Child Develpm.*, 1944, 15, 145-166.
5. WENAR, C. The effects of a motor handicap on personality: I. The effects on level of aspiration. *Child Develpm.*, 1953, 24, 123-130.
6. WENAR, C. The effects of a motor handicap on personality: II. The effects on integrative ability. *Child Develpm.*, 1954, 25, 287-294.

A FACTOR ANALYSIS OF SOME ATTENTION-SEEKING BEHAVIORS OF YOUNG CHILDREN[1]

Jacob L. Gewirtz[2]
The University of Chicago

Problem

Though knowledge of the process underlying young children's behavior for attention and similar reinforcers might well be a key to understanding their general social interaction with both adults and peers, very little is known either about the dimensionality of their attention-seeking or about its antecedents. In the study herein reported an attempt is made to gain some leverage on the dimensionality of the behavior classes employed for attention by young children through a factor analysis of the correlations among nine items of behavior judged to be consonant with the characteristics generally attributed to attention-seeking and age.

Behavior for the reinforcer provided by the attention of people (attention-seeking) has usually been defined as falling under the more broadly defined concept "emotional dependence," along with such behaviors as those employed for approval, praise, physical contact and caresses, reassurance, and nearness. While the clinical literature is replete with case histories, mostly based on informal observation of children and adults with adjustment problems, which are described and interpreted in terms of such behavior concepts, the research literature reveals that few systematic attempts have been made to study the dimensionality of either emotional dependence or attention-seeking, or to determine their antecedents (1, 2, 3, 8, 12). Typically, relationships involving emotional dependence and attention-seeking in children have been found as by-products of general studies which addressed themselves to the consequences in children's behavior of broadly defined family attitudes or conditions of socialization, such as overprotection, deprivation, or rejection (e.g., 5, 6, 7, 10).

In order to further a systematic experimental investigation of the behaviors employed for attention and the antecedents of those behaviors, an experiment (4) was conducted on young children, which employed a situation in which each child could easel paint as long as he wished in the presence of an adult. The purpose was to investigate within the situational

---

[1] This investigation was supported by a grant-in-aid from the Social Science Research Committee of the University of Chicago.

[2] The writer is grateful to Dr. L. V. Jones for his advice with respect to the factor analysis.

CHILD DEVELOPMENT, Vol. 27, No. 1 (March, 1956)

context of the experiment the relationships obtaining among several classes of behavior which logically could be expected to have been employed for attention (notice) from the adult, and three independent variables: the degree of apparent availability of the adult to the child (low versus high), the sex of the adult, and the sex of the child. The assumption underlying the use of these independent variables was that they provided stimulus conditions which were related to critical socialization experiences of the children, and therefore that singly or in combination they might be critical determinants of the behaviors of young children for the attention and perhaps the approval of adults. (Although attention- and approval-seeking in principle can be distinguished, for example by demonstrating differential emission of these behaviors as a function of whether attention or approval is delivered by the adult, for purposes of this investigation the two modes of behavior were formally confounded in the interests of the probable generality of the resulting data.)

The six response variables employed in the experiment were scored on a basis of frequency of occurrence. It was assumed that the first three variables listed very likely would be employed for the adult's attention (and/or approval) and that the last three variables also might be employed for that reinforcer. The variables were: *Attention-seeking,* which included urgent responses for overt notice from the adult; *Questions* casually asked, which required only brief replies from the adult; *Comments* casually made, which required no response from the adult; *Glances* directed momentarily toward the adult; number of *Paintings* completed; and *Time* in minutes the child remained in the session.

It was discovered that the children evidenced reliably greater mean frequencies of Attention-seeking and Glances under the low than under the high adult availability condition. And although, for the two conditions of availability, the means of Questions, Comments, Paintings and Time rank ordered identically as did the means of Attention-seeking and Glances, they were not reliably different from one another.

It was found also that the boys evidenced a reliably greater mean frequency of Attention-seeking responses with women than with men, but that the difference between the greater mean frequency of such responses employed with men than with women by the girls was not reliable. This differential response pattern was duplicated by the mean frequencies of Comments and Questions.

However, the dimensionality of attention-seeking, to the extent that it might be revealed in this experiment, is but incompletely indicated by the differential relationships between the dependent variables and the independent variables, and could be clarified by an analysis of the relationships among the six response variables employed. And since data bearing upon three additional response variables of the same class, *Help-seeking, Permission-seeking,* and *Praise-seeking,* which were not included in the above study, were available, as was *Age* (in months) of the subjects (which could serve as a useful reference variable), the 10 variables were intercorrelated. A factor analysis of this inter-item correlation matrix could be of considerable use in determining the dimensionality of attention-seeking (and perhaps approval-seeking as well), especially when considered with the findings of the experiment, despite the limitations imposed by the moderate proportions of the matrix, the questionable or unknown nature of some of the distributions, observer reliabilities, and the stability through time of some of the behavior items, and the moderate size of the sample employed. Moreover, the results of the factor analysis could perhaps aid in interpreting the results of the experiment. The interpretations of the factor analysis are reported in this paper.

Experimental Procedure

Since the experimental method which generated the response variables analyzed is reported in detail elsewhere (4), only its critical features will be summarized here.

Sample

Twenty-eight boys and 28 girls were selected at random from among the white middle-class children at the University of Chicago Nursery School, with one control exercised: the age distributions were made relatively homogeneous for the eight experimental conditions. The subjects ranged in age from 4 years, 0 months, to 5 years, 7 months. The median age of the 28 boys was 4 years, 8 months; of the 28 girls, 4 years, 7 months.

Procedure

Each subject was escorted into the experimental room by the adult and informed that he could easel-paint as long as he wished, that he could request sheets of newsprint upon which to paint as he required them, and that when he no longer wished to paint, he was to inform the adult who would escort him back to his play group.

Independent Variables

The three experimental variables, degree of availability of the adult to the child, sex of the adult, and sex of the child, each included two levels in a 2 x 2 x 2 factorial design. Each child participated once under only one of the eight conditions of the experiment, making for eight treatment groups of seven children each.

Degree of adult availability. The two availability conditions differed only in one characteristic of the adult's role, his apparent availability to the child. Under the *high* availability condition, the adult sat close to the child, giving complete attention to him. Under the *low* availability condition, the adult sat at a somewhat distant desk, busily working on papers. The adult responded to the initiations of the child in a uniform, standard manner under both conditions. His response was always friendly, yet brief, tailored to the character of the child's initiation, and engineered to terminate the interaction sequence. Responsibility for continuation of the sequence was always the child's. Thus the adult was as receptive to the children under the low as under the high availability condition.

Sex of adult interactor. Two men and two women, familiar to the children, each served as interactor under the four conditions which involved the adult of his or her own sex. Each served with approximately one-half of the subjects under every one of those four conditions.

Sex of child subject. Of the eight cells of the experimental design, four each included seven boys and four each included seven girls.

Dependent Variables

Of the 10 variables whose intercorrelations were factor analyzed, the relationships to the independent variables of the first six listed were studied in the experiment reported. It was postulated that under the conditions of the experiment each of the response classes might logically be expected to have been employed for the reinforcer provided by the receipt of attention (or approval) from the adult.

Attention-seeking. Included were responses, characterized by urgency, apparently designed for active notice from the adult. The adult's overt response indicated, in a manner appropriate to the child's request, that he was attending to the child. If an immediate response from the adult was not forthcoming (the adult delayed his response for three seconds, to insure correct categorization of the child's behavior), possibly the child would repeat his response several times until noticed. A typical sequence involved the child making his demand or asking his question, then pausing to direct his complete attention toward the adult while awaiting his reply. After the adult had responded, the child relaxed, either to continue his painting or to originate another sequence of behavior. Included in this category were requests and demands to have a statement heard, and responses like the following: "Will you look at this picture?" "What do you think this is a picture of?" "Look at me!" "You know what?"

Questions. Included in this category were casual questions, never urgent, which required only brief replies from the adult. The child typically continued his painting while awaiting a reply, and if he looked toward the adult at all, he did so momentarily. Aside from the fact that a reply from the adult was usually required, the tenor of the child's behavior was very much like that characteristic of Comments. An example is, "Why is the window open?"

Comments. Included were casual remarks (never urgent), which usually required no response from the adult. Typically the child made his remark, and without pause either continued on to another remark or resumed his painting. He never attempted to insure either that the adult would hear or respond to his remark.

Glances. Included were the child's glances and stares at the adult, which neither involved verbal accompaniment, required responses from the adult, nor initiated sequences of interaction when it was apparent to the child that his glance had been noticed by the adult.

Paintings. When a child had informed the adult that he had completed a painting, the adult removed the painting from the easel, put up a blank sheet of newsprint in its place, and wrote the child's name in a corner of that sheet. The number of paintings on which the child worked during the session was recorded.

Time. The time elapsing from the point at which the subject took his place at the easel until he stated that he no longer wished to paint was recorded to the nearest minute.

Behaviors falling under the three categories listed below were recorded also, though they were not employed in the experiment:

Praise-seeking. Included were requests or demands for praise from the adult. In every other respect, the characteristics of this behavior class were identical with those of Attention-seeking (e.g., "Isn't my picture pretty?").

Help-seeking. Included were requests or demands for instrumental help or guidance from the adult (e.g., "Help me take this picture down").

Permission-seeking. Included were requests for permission from the adult for performance of an act (e.g., "What am I allowed to paint?").

Age. A tenth dependent variable, included to provide reference, was age in months of the subjects.

Reliabilities of the Response Variables

The observer reliabilities of the behavior categories, presented in the form of per cent agreement, were determined from two series of simultaneous yet independent observations made in the experimental setting by three observers. In the first series, 12 subjects not included in the sample were observed; in the second series, 16 of the first 30 experimental subjects were observed. Each of the eight experimental treatments was sampled at least once. The subject's responses were categorized both according to the definitions of the behavior variables and a standard recording procedure.

Every observer's record was aligned with the record of every other observer for each subject observed. All the records for each pairing of observers were combined for each observation series. Per cent observer agreement was computed according to the usual formula:

$$\% \text{ Agreement} = [\# \text{ Agreements} / (\# \text{ Agreements} + \# \text{ Disagreements})] \times 100.$$

An agreement was indicated by the assignment of a particular behavior to the same behavior category by two paired observers, and a disagreement was

EDNL. & PSYL. 1-9-70 LIBRARY

indicated when they assigned a particular behavior to different categories or when one observer failed entirely to categorize the behavior.

The median of the per cent agreement figures of the two observation series was computed for each of the three observer pairs on each behavior variable; and the median of these three medians was computed in turn for each variable. The per cent agreement figures follow: Comments, 76; Questions, 80; Attention-seeking, 77; and Glances, 82. Because of the moderate occurrence of Help-seeking during the reliability observations, the observations of the three observer pairs were combined, to produce the single agreement figure, 73 per cent. The incidence of both Permission-seeking and Praise-seeking was low during the reliability observations and precluded obtaining meaningful reliability estimates for them.

The observer reliability figures for the behavior variables Comments, Questions, Attention-seeking, and Glances appear adequate; and while of questionable validity the reliability figure for Help-seeking is reassuring. The reliabilities of Permission-seeking and Praise-seeking are unknown. (It is probable that the low frequency of occurrence of Help-seeking, Permission-seeking and Praise-seeking under the conditions of the experiment, which precludes the effective assessment of their observer reliabilities, accounts also for their questionable distributions in the experiment.) Response

Table I

THE INTER-ITEM CORRELATIONS *

	Attention-seeking	*Questions*	*Comments*	*Glances*	*Paintings*	*Time*	*Age*	*Praise-seeking*	*Help-seeking*	*Permission-seeking*
Attention-seeking	...	.631	.770	.235	.406	.241	.064	.500	.549	.398
Questions		...	.684	.122	.373	.010	—.146	.463	.571	.403
Comments			...	.213	.465	.253	.041	.415	.477	.331
Glances				...	.416	.589	—.110	.144	.086	.208
Paintings					...	.713	.277	.236	.404	—.009
Time						...	.153	.193	.304	.171
Age							...	—.029	.087	—.227
Praise-seeking								...	.368	.315
Help-seeking									...	.224
Permission-seeking										...

Note.—With the exception of Age, the scores for all variables were transformed into logarithms before correlation.

* The coefficients which are significant at the different levels, for 54 *df*, are: $p < .001$, .43; $p < .01$, .34; $p < .05$, .26; $p < .10$, .22.

category consistency, based on repeated observations of the subjects through time, was not determined.

The Correlations

In order to lessen the departure from normality of the distributions of frequency scores and to stabilize their variances, logarithmic scores were employed for the nine behavior variables. The Age (in months) variable was not transformed. Product-moment coefficients were calculated for the correlations among the 10 variables.

Because the frequency distributions of the variables Help-seeking, Permission-seeking, and Praise-seeking included relatively many zero scores, it is doubtful that they would be normalized effectively by transformation. Thus the correlation coefficients between these three and the other seven variables are not readily interpretable. And since their observer reliabilities are for all purposes unknown, the coefficients involving the three variables have been segregated and listed last in the table of inter-item correlations (Table 1), and the results of this study which depend upon them will be interpreted with considerable caution.

The correlation coefficients employed were calculated from the transformed experimental scores. Since the experimental procedure and the treatment employed in the study have been established to represent conditions under which children do make those initiations which fall under the nine behavior item categories employed, and since several of these categories were successfully related to independent variables hypothesized to be relevant, it would not seem unreasonable to generalize the correlations obtained to the variety of social settings in which a child and an adult are involved.[3]

THE FACTOR ANALYSIS

Method

The centroid method of multiple factor analysis, developed by Thurstone (11), is employed in this study. The matrix of centroid factor loadings is presented in Table 2. Four centroid factors were extracted from the correlation matrix.

In the present investigation, the decision was made to call centroid factor IV a residual for three overlapping reasons: first, seven of its 10 factor loadings were of zero order; second, it accounted for so little of the

---

[3] An analysis of the differences between the correlation coefficients employed (54 *df*) which were determined upon the experimental scores, and the within-groups (residual, error) coefficients (47 *df*) for the intercorrelations among the first seven variables in Table 1 can support this point. Within-groups correlations are by their natures corrected for the effects of experimental treatments. In principle, then, they might be generalized to a wide variety of treatment conditions. The analysis indicates that for the 21 pairs of coefficients involved in the comparison of experimental score correlations with the within-groups correlations, the median difference between members of pairs was .02, and the largest difference .11. In no case is it likely that a difference is statistically stable.

TABLE 2

CENTROID FACTOR LOADINGS

Item	Centroid Factors I	II	III	IV	h^2 *
1. Attention-seeking	.816	.301	.058	.261	.828
2. Questions	.693	.508	.102	—.308	.844
3. Comments	.774	.245	.156	.190	.720
4. Glances	.421	—.339	—.402	.036	.455
5. Paintings	.729	—.512	.255	—.101	.869
6. Time	.607	—.678	—.238	.023	.885
7. Age	.054	—.236	.384	.052	.209
8. Praise-seeking	.525	.232	—.082	—.019	.337
9. Help-seeking	.627	.144	.147	—.119	.450
10. Permission-seeking	.376	.335	—.380	.037	.399
$\Sigma k^2/n$ †	.364	.149	.064	.023	.600‡

* Communality.

† Estimate of proportion of total item variance attributable to a factor.

‡ $\Sigma h^2/n = \Sigma(\Sigma k^2/n)$.

average item variance that it appeared highly improbable that it would yield an interpretable factor; and, third, it did not yield simple structure upon rotation.

Rotations. Orthogonal rotations of the centroid factors were accomplished employing a graphical modification of Thurstone's methods (13). An orthogonal solution rather than an oblique solution was employed primarily because of its advantages of relative conceptual simplicity and economy of labor. It is not assumed that if the oblique method were employed, the factors extracted necessarily would be independent of one another. In

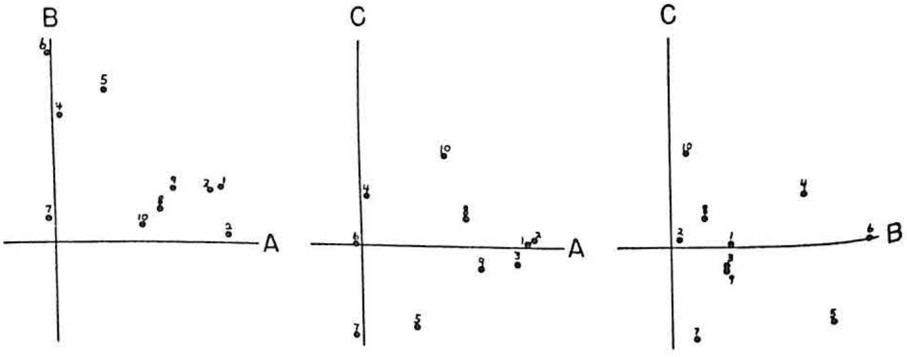


FIGURE 1—Two-dimensional plots of each rotated factor against the other two factors.

TABLE 3

ROTATED FACTOR LOADINGS

Item	*Rotated Factors* A	B	C	h^2 *
1. Attention-seeking	.82	.28	.02	.75
2. Questions	.85	.04	.04	.73
3. Comments	.77	.27	—.09	.67
4. Glances	.03	.64	.25	.47
5. Paintings	.27	.77	—.38	.81
6. Time	—.02	.95	.00	.90
7. Age	—.03	.12	—.45	.22
8. Praise-seeking	.52	.17	.14	.32
9. Help-seeking	.58	.27	—.11	.42
10. Permission-seeking	.42	.08	.46	.39
$\Sigma k^2/n$ †	.28	.22	.07	.57‡

* Communality.

† Estimate of proportion of total item variance attributable to a factor.

‡ $\Sigma h^2/n = \Sigma(\Sigma k^2/n)$.

Table 3 appear the rotated factor loadings, while in Figure 1 appear the two-dimensional plots of each rotated factor against the other two factors.

Criteria for interpretation of factor loadings. Because of the questionable validity of the correlations involving the three variables Praise-, Help-, and Permission-seeking, the interpretation of a factor defined only by one or more of those variables would be seriously limited indeed. For this reason, and because of the moderate number of cases studied, rather conservative criteria are employed in the interpretations of the factor loadings, as follows: Loadings of ±.60 and above are considered high, and are taken to define the meaning of a factor; loadings of ±.30 to .59 are considered moderate, and must be consistent with the definition of the factor; loadings of ±.20 to .29 are considered low, and while not taken seriously into account, they are employed to clarify the meaning of a factor; and loadings of less than ±.20 are considered negligible.

We may now move to the consideration of the interpretations of the factors.

Factor A

The factor most clearly identified in the present study is Factor A. In Table 3 it is seen that this factor accounts for the largest proportion (.28) of the total item variance. In that seven out of the nine behavior items show appreciable loadings on A, it might almost be called a general factor. Since the communalities of the items for the three rotated factors average .57, Factor A accounts for approximately 50 per cent of the variance in the item intercorrelations accounted for by the three factors extracted.

Those behavior items having projections on Factor A, which are of interest according to the criteria advanced, are listed below:

Item	*Loading on Factor A*
2. Questions	.85
1. Attention-seeking	.82
3. Comments	.77
9. Help-seeking	.58
8. Praise-seeking	.52
10. Permission-seeking	.42
5. Paintings	.27

The three behavior items which define this factor are verbal in character. Thus, the factor might ordinarily be labeled "talkativeness" or the like. However, in this experiment, where these behavior items related differentially to the experimental variables roughly according to hypotheses advanced, it appears reasonable to emphasize that characteristic of them which was responsible for their initial selection. In the context of the experimental setting, Attention-seeking, Questions, and Comments all appear to involve an active attempt *to gain or to maintain the adult's attention*. The major dimension along which the behavior items appear to differ is in the degree to which their initiations for attention are overt or direct. The analysis, which is detailed below, indicates that this interpretation of the factor solution fits well with a number of reasonable assumptions.

Attention-seeking responses were apparently employed by the child overtly to gain the attention of the adult; he paused in his ongoing action, giving his complete attention to the adult until his request had been answered by the adult's overt response.

Casual Questions overlapped with Attention-seeking in two ways. First, urgent questions have been defined in this study to fall under the Attention-seeking behavior category; on this basis alone Questions might have been expected to enter into relationships parallel to those found for Attention-seeking, and, perhaps as well, relationships with Attention-seeking. And, second, while it was not as apparent as was the case with Attention-seeking that Questions were employed to gain the attention of the adult, since they might have been employed as well for other reinforcers (such as those provided by the information included in the overt replies they required), the demand for a reply which is inherent in the nature of the casual question actually could gain the attention of the adult for the child. Thus, Questions may frequently have been employed by the children solely for the reinforcer provided by the adult's attention.

While casual Comments did not appear overtly to have been employed to gain the adult's attention, in a room of moderate size where a distance of no more than six feet separated the child from the permissive adult, it would not have been unreasonable for the child to expect that his every remark would be noticed by the adult. Thus, while Comments, like Questions, might have been emitted for reinforcers other than those provided by the adult's attention, because of the nature of the interaction setting, they reasonably might be expected to have been employed frequently to gain that attention.

The interpretation of Factor A advanced in the preceding paragraphs would apply as well to Help-seeking, Praise-seeking, and Permission-seeking, which three behavior items have moderate loadings on the factor. Each of those behaviors could also be expected to have been employed to gain or to maintain the attention or approval of the adult.

Help-seeking logically might have been employed solely for the reinforcers involved in the receipt of help. However, since such requests required an overt response from the adult, they could effectively have gained his attention. Moreover, requests for help could place the child in a dependent role relative to the adult; and since it could have been expected that an adult would ordinarily offer attention to a dependent child, some children might have attempted to appear dependent in order to gain the adult's attention.

Praise-seeking might be expected to have been employed to gain the approval (and attention) of the adult, and is seen to have much in common with Attention-seeking. Similarly Permission-seeking could be expected to have been employed to retain or to gain the approval of the adult, since by such requests children can appear compliant ("good"). Moreover, since Permission-seeking requests require some response from the adult, they might have been employed on occasion directly for the attention provided by the adult.

It is noted in concluding this interpretation that not one of the verbal behavior items is missing from the list of items which have at least moderate loadings on this factor. Moreover, it is reassuring to note that the three behavior items about which some question was raised have moderate loadings on this factor, and are quite consistent with the interpretation advanced.

Factor B

Next most clearly identified in the present study is Factor B, whose loadings account for 22 per cent of the average item variance. On this factor, three items have high loadings, while three items have low to moderate loadings. The factor accounts for approximately 37 per cent of the variance in the inter-item correlations accounted for by the three factors extracted. The items which have appreciable loadings on Factor B are:

Item	*Loading on Factor B*
6. Time	.95
5. Paintings	.77
4. Glances	.64
1. Attention-seeking	.28
3. Comments	.27
9. Help-seeking	.27

Factor B describes a pattern wherein children who remain in the experimental setting for longer periods of time completed greater numbers of paintings and employed greater numbers of glances at the adult. Put somewhat differently, the children who remained in the setting longer, and completed more paintings, emitted relatively few verbal responses to the adult (identified as attention-seeking according to the interpretation of Factor A). Rather they glanced at the adult frequently.

The three behavior items which define Factor B are nonverbal in character. Both Time, which is essentially a pure item factorially, and Paintings are characterized by compliance with the instructional set given the children to paint. Several alternative and more or less reasonable interpretations of Factor B can be advanced on the basis of the three behavior items which define it. The factor might involve a response pattern which:

a. Reflected the child's enjoyment of the painting task, and his consequent expenditure of time in it.

b. Reflected passive compliance with the requirements of the instructional set by the disinterested, docile child, who might perhaps have been socially apprehensive as well.

c. Was employed for the adult's attention and approval by the passive, docile child who hesitated to employ active verbal responses like those which characterize Factor A; or by the child who for some reason was in conflict over employing that more direct response pattern.

d. Was achievement-oriented to the task prescribed for the adult's attention and/or approval. This pattern might have been employed by relatively mature children, for whom compliance was the most appropriate response for the adult's attention or approval, and/or by the children who did not "seek" verbal assurance that the adult was available to them.

Of the four possibilities, interpretations *c* and *d* are believed to be the most reasonable, although interpretations *a* and *b* present the greater appeal to apparent parsimony. This conclusion is supported by the following reasoning: Both the child deriving considerable enjoyment from the task itself and the compliant, docile child disinterested in the task might have evidenced high scores on Time, Paintings, and Glances. These types could reasonably be expected to spend a considerable amount of time in the situation, either because of enjoyment of the task, or to satisfy the adult's instructions. On either basis, it follows that a relatively large number of paintings would be completed. And finally, if it is assumed that Glances have a reasonably uniform distribution through time, i.e., that their rate is fairly stable, then expenditure of time implies a correspondingly high frequency of Glances.

However, in the analysis of the results of the experiment, when the relationship between Glances and the independent variable Degree of Adult Availability was assessed and the glance scores corrected for their regression upon Time, it was found that children employed reliably greater frequencies of both Attention-seeking and Glances under low availability than under high availability, although the correlation between those two dependent variables did not attain significance at the .05 level. Thus, Glances could not have been accounted for completely by Time, nor could the relationship between the two have been accounted for by the independent variable Degree of Adult Availability. Hence, it follows that Glances reflects other determinants, presumably related to attention- or approval-seeking, in addition to those reflected by Time, and this finding detracts seriously from the apparent parsimony of interpretation *a*, and, to a lesser extent, of interpretation *b*. Put differently, it appears that there might have been at least two (perhaps more) qualitatively different types of glances.

Glances might have been employed directly by the child for the adult's attention and approval, for compliance, for painting products, or for "being a child" in the situation; for the child might well have expected that if his glance were noticed by the adult, he would be attended to. As such, Glances might have been responses alternative to Attention-seeking, Questions, and Comments, in that a high glance score for a child might mean only that on a high proportion of times he looked at the adult, he made no verbal requests or statements.

In addition to the possibility that they were employed for attention and approval, Glances might have been employed by the children simply to help them to ascertain what was occurring around them, or pehaps to understand the situation which was not, after all, one with which they were highly familiar. Moreover, Glances might have represented simply attempts by children to ascertain if the adult was still content to allow them to continue with the enjoyable painting task. And it is also possible that Glances might simply have been indicators of uneasiness or apprehension in the children.

If the loading of Glances on Factor C, which will be interpreted provisionally to involve inexperience or apprehensiveness, is taken to reflect the proportion of occasions when they served as indicators of uneasiness, then their loading on Factor B might be taken to reflect primarily behavior for the reinforcers of attention or approval alone. This logic would lend support primarily to alternative factor interpretations *c* or *d*.

These interpretations receive additional support from the following reasoning: Since the adult had informed the child that he might paint as long as he wished, the child must have realized that to remain in the setting and to paint would satisfy the adult. Moreover, it is recalled that when the child informed the adult that he had completed a painting, the adult took it down from the easel, replaced it with a fresh sheet of newsprint and wrote the child's name in the corner of that sheet. In servicing the easel, then, the adult attended to the child, and to the extent that completing a painting was an achievement, the adult's attention would have signified approval to the child. Thus, it is in order to suggest that the child might have completed paintings to gain the attention and approval of the adult, as it may be suggested that the response of remaining in the setting, reflected in Time, might have been employed in a like manner.

The remaining factor loadings of Attention-seeking, Comments, and Help-seeking, all low, would appear to be in keeping with interpretations *c* and *d*. Attention-seeking and Comments define and Help-seeking has a moderate loading on Factor A, which was interpreted to involve active verbal attention-seeking.

Thus, though it is not possible from the pattern of loadings found for this factor conclusively to select from among the several possible interpretations, it is thought most likely that Factor B reflects behavior for the reinforcer of attention (approval), and that it does not reflect inexperience, apprehensiveness, or simple enjoyment of the task. However, it is conceivable that several of the interpretations advanced hold simultaneously; for several of the hypothesized types of children—active approval-oriented, docile approval-oriented, docile disinterested, and those who simply enjoyed painting—characterized by this factor pattern might have been represented in the sample employed.

In order to allow selection from the possibilities discussed, a factor analysis study could be planned which would include, in addition to the variables employed in the study, response variables selected effectively to isolate the different interpretations. Alternatively, selected independent variable treatment conditions could be employed to serve a similar end.

Factor C

Least clearly identified in the present study is Factor C, which accounts for only 7 per cent of the average item variance. While no item is found to have a high loading on this factor, three items have moderate loadings and one item has a low loading. The items which have non-zero loadings on this factor are listed below:

Item	*Loading on Factor C*
10. Permission-seeking	.46
7. Age	—.45
5. Paintings	—.38
4. Glances	.25

Ordinarily one would hesitate to attempt to interpret a factor which is specified by items with only moderate loadings, and one would be especially cautious when the factor accounts for such a small amount of the average item variance. However, in this instance the meaningful grouping of the items would appear at least to support an attempt at interpretation.

An examination of the items with moderate loadings on this factor indicates, although clear definition is not possible, that the tendency embodied may be interpreted provisionally as involving either (or both) of two related processes, both perhaps to some extent determined by the relative immaturity of children in the age range from four to five and one-half years: (*a*) apprehensiveness, either general in nature or about losing the adult's approval; and/or (*b*) inability or inexperience in painting.

It is recalled that when Factor A was interpreted, the suggestion was advanced that requests for permission might be expected to have been employed to gain the adult's attention; and, in that they entailed compliance, they might have been employed to retain the adult's approval. It is reasonable to expect that young children in the age range studied, being relatively immature, would tend to employ such responses. Moreover, younger children might find it difficult to concentrate on producing paintings, or not have the ability to do so. And following from the assumption that the high loading of Glances on Factor B accounts for their having been employed for attention and approval, their loading on Factor C might reflect their having been indicators of inexperience and apprehensiveness.

Thus, it appears that the behavior pattern wherein younger children request permission more frequently, complete fewer paintings and emit more frequent glances, makes considerable sense. However, it is again proper to emphasize that this factor was not most clearly defined, that Permission-seeking has both a questionable distribution of scores and an essentially unknown reliability, that it is impossible to select from among the alternative interpretations (which overlap considerably), and that other reasonable interpretations might be possible.

DISCUSSION

The basic logic of the approach of this study has been that the behaviors scored could be said to have been employed for the reinforcers of attention, and possibly approval, on a basis of the structure of the experimental situation. That is, on certain logical and empirical grounds a situation was devised in which the emission of certain behaviors (which in other contexts might mean a number of things) could logically be expected to represent attention-seeking; and those behaviors were defined as such. Since the design of this study precluded the demonstration of the existence of a functional relationship between the emission of these behaviors and the receipt of attention, the logic of the nature of the experimental setting stands as the basic justification for labeling the behaviors studied attention-seeking; and the consonance of the factor loadings with this interpretation and the reassuring relationships obtaining between several of these dependent variables and certain combinations of experimental variables constitute auxiliary support for the basic assumption employed (and constitute as well new information about the behaviors).

The Results

On the basis of an analysis of their characteristics it was expected that some of these behavior items might fall together to define factors. In general this proved to be the case, particularly for the six behavior items whose relationships to three experimental variables were determined in the investigation reported earlier. It was assumed for the experimental conditions employed that Attention-seeking, Questions, and Comments might all three be employed similarly for the reinforcer provided by the attention of the adult, and therefore that they might correlate with each other and relate similarly to the independent variables. The results of the factor analysis to an extent support that assumption, and indicated that those three behavior items all have rather high loadings on, and defined, Factor A, which was interpreted provisionally to involve active attempts to gain the attention of the adult. It will be recalled that although the means for those three variables tended to rank-order similarly for the different experimental combinations, reliable effects were found through an analysis of variance only for Attention-seeking, while the effects for Comments approached significance.

It was assumed further for the conditions of the experiment that the nonverbal behavior items, Paintings, Time and Glances, might also have been employed for the reinforcer provided by the adult's attention or approval. It was expected, therefore, that they might relate to Attention-seeking, Questions and Comments, and enter as well into similar relationships with the experimental variables, or that they might in both cases relate as might an alternative response pattern.

It will be recalled that one of the alternative interpretations of Factor B was that it involved a nonverbal pattern of behavior for the reinforcer of attention or approval, which was different from the active verbal form characteristic of Factor A. While Time, Paintings, and Glances had high loadings on Factor B, with the exception of the low loading of Paintings, they had no loadings on Factor A, as should have been the case if they were highly related to Attention-seeking, Questions and Comments. Moreover, neither Time nor Paintings related at all to the experimental variables. Glances however related to the apparent availability conditions similarly as did Attention-seeking. The ambiguity involved in the interpretation of Factor B does not detract from its well-defined status in this investigation.

Because the variables included in the analysis of variance differed in several important respects from those included in the factor analysis, it would not be fruitful to compare in any greater detail the results of the experiment, in which the behavior items were related to independent variables, with the results of the factor analysis, in which the common elements in the relationships among the behavior items were determined.

It is of interest to note that the nine behavior items selected initially in the context of the experiment to appeal to a presumed single dimension of attention-seeking, should (with the reference variable of age) yield at least two and perhaps three factors. This apparent multidimensional character of attention-seeking in experimental settings like the one in this study could provide a useful basis for future investigations of both its dimensionality and its antecedents.

Behavior Antecedents

The results obtained indicate that it was not unreasonable to advance as a basic assumption the proposition that behaviors like Attention-seeking, Questions, and Comments directed toward adults were frequently reinforced (particularly with generalized reinforcers (9) like attention and approval) during early childhood, and hence that an experimental setting which furnished children with access to a permissive adult would provide the occasion for such behaviors to be employed for those reinforcers.

It is assumed that for any child the hierarchy of reinforcers and of responses employed to gain those reinforcers would depend upon his particular reinforcement history. Thus, some children might characteristically seek approval, while others might simply seek attention; and some children might characteristically employ behavior overtly directed toward gaining a particular reinforcer (e.g., Attention-seeking), while others might employ behavior which perhaps could gain the reinforcer indirectly (e.g., Questions or Comments).

It is conceivable that other than simple characteristics of the reinforcement histories of the children could account also for the different degrees to which different responses (e.g., those which might be characterized as either direct of indirect) are employed for attention. If it is assumed that the overt (direct) response for attention most readily gains that reinforcer, then it is conceivable that those children who might have been in conflict over employing that response (which can be assumed to be a joint function of the subject's history of positive and negative reinforcement for the response) might have employed a displaced (e.g., indirect) response for attention. Of course, the responses labeled indirect might have been employed frequently because they have been learned as appropriate in given circumstances, and not as responses to conflict.

Moreover, the indirect response for attention might have been employed by children on occasions when, for whatever reason, it was critical for them to gain attention, for the indirect response by its nature does not appear readily to invite rejection, and could have insured that an adult would take notice of the child. Likewise, characteristics of the sequence of responses might have played a role in determining whether, for example, direct or indirect responses were employed for attention. When a response over which the child was in conflict was employed for any reason, the child's conflict might have become heightened, and as a consequence, subsequent responses emitted by him in the series might have been displaced.

If it is assumed that a process like that of conflict might have operated at least in part to determine the relationships among the behavior variables in this investigation, then those relationships would have been obscured. An investigation which introduced differential conflict in terms of selected independent variables would be required to gain leverage upon the question.

In concluding these considerations, it appears not unreasonable to assume that the generalized reinforcer (9, p. 77) provided by the attention or by the approval of the adult, which so frequently determines the child's behavior, is effective because it signifies to the child (on some latent level), that upon further request he might receive other, critical reinforcers (food, etc.) which an attending or approving adult can supply.

Methodological Considerations

It appears that the factor analysis has helped to organize a domain which heretofore has been essentially unexplored. Since this study opens a new domain to investigation and there is little precedent which can be employed to aid in the interpretation of the findings, reason for caution obtains. Rather than providing decisive evidence for a given hypothesis, the interpretations of the results are indicative of the directions in which future investigations might move fruitfully. The factors isolated suggest how the behavior items employed might be combined or improved, and, as well, the natures of other behavior items which might be investigated.

Although a number of the apparently reasonable factor interpretations are consonant with assumptions of attention- (and approval-) seeking, it is well to emphasize that all the factor interpretations advanced in this paper are *ad hoc*; and further, that they are limited by the subjective element in factor rotations, by sampling instability of the inter-item correlation coefficients, by unknown observer reliability and item consistency of several of the behavior items, and by the fact that the data were collected with reference to the treatment conditions of an experiment.

In the present study factors were defined by considering the nature of items exhibiting high factor saturations. As a task for the future, a more comprehensive analysis would be desirable, in which a larger number of reliable behavior items would be employed, perhaps selected to differentiate among the alternative factor interpretations, as would be a larger number of subjects to insure sampling stability, and behavior items which yield more normal distributions. The nature of the items could be more adequately determined in this manner, and factors could be defined with greater confidence. Moreover, by employing as subjects children at several age levels, the developmental characteristics of the factors could be explored.

SUMMARY

In an experiment reported elsewhere (4), each of 56 children could easel paint as long as he wished in the presence of an adult. The experimental

conditions employed were degree of apparent availability of the adult to the child (low versus high), sex of the adult, and sex of the child. The children's behaviors were classified in nine categories, each of which, according to reasonable assumptions, could be expected to have been employed in the experimental context for the reinforcer of attention or approval from the adult. It was thought that a factor analysis of the inter-item correlation matrix could be of use in clarifying the dimensionality of attention-seeking.

The nine behavior items, together with age, were intercorrelated, forming a matrix of moderate proportions. This matrix of intercorrelations was subjected to a graphical modification of Thurstone's methods of multiple factor analysis. After orthogonal rotations of the four centroid factors which were extracted, three apparently meaningful factors appeared. Factors A and B were more clearly defined than Factor C. However, only Factor A, and to a lesser extent Factor C, appeared to lend themselves to ready psychological interpretations. The factor interpretations follow:

Factor A which was most clearly defined was interpreted as involving an active (verbal) attempt to gain or to maintain the attention of the adult.

Factor B, which was next most clearly defined, was difficult to interpret. Several alternative and more or less reasonable interpretations were advanced. It was thought most likely that the factor involved either: (*a*) an achievement-oriented response pattern to the task prescribed, employed for the adult's attention or approval by relatively mature children or by children who did not require verbal assurance that the adult was available to them; or (*b*) a response pattern for attention or approval employed by the passive, docile, child, or by the child in conflict over employing the more direct response pattern. That Factor B reflected either the child's enjoyment of the painting task or passive compliance with the instructions by the disinterested docile child could not be ruled out completely as interpretations.

Factor C, which was least clearly defined, was interpreted provisionally as involving either or both of two related processes, perhaps to some extent reflecting relative immaturity of some of the children, as follows: apprehensiveness, either general in nature or about losing the adult's approval; and/or inability or inexperience in painting.

Assumptions were set forth concerning the role of certain stimulus events in critical socialization settings, and implications for consequences susceptible to experimental investigation were discussed.

Because of limitations in the data and the method, the results were evaluated with considerable caution. It was concluded that the factor analysis helped, if only to a modest degree, to organize a domain which heretofore has been essentially unexplored.

REFERENCES

1. BELLER, E. K. Dependency and independence in young children. *Amer. Psychologist,* 1950, 5, 293. (Abstract)

2. Gewirtz, J. L. Succorance in young children. Unpublished doctor's dissertation, State Univer. of Iowa, 1948.
3. Gewirtz, J. L. An investigation of nurturant behavior in young children. *Amer. Psychologist,* 1949, 4, 248. (Abstract)
4. Gewirtz, J. L. Three determinants of attention-seeking in young children. *Mongr. Soc. Res. Child Develpm.*, in press.
5. Goldfarb, W. Infant rearing and problem behavior. *Amer. J. Orthopsychiat.,* 1943, 13, 249-265.
6. Hattwick, B. W. Interrelations between the preschool child's behavior and certain factors in the home. *Child Develpm.,* 1936, 7, 200-226.
7. Levy, D. *Maternal overprotection.* New York: Columbia Univer. Press, 1943.
8. Sears, R. R., *et al.* Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.,* 1953, 47, 135-234.
9. Skinner, B. F. *Science and human behavior.* New York: Macmillan, 1953.
10. Symonds, P. M. *The psychology of parent-child relationships.* New York: Appleton-Century, 1939.
11. Thurstone, L. L. *Multiple-factor analysis.* Chicago: Univer. of Chicago Press, 1947.
12. Whiting, J. W. M., & Child, I. L. *Child training and personality* New Haven: Yale Univer. Press, 1953.
13. Zimmerman, W. S. A simple graphical method for orthogonal rotation of axes. *Psychometrika,* 1946, 11, 51-56.

THE EFFECTS OF NUMBER OF REINFORCEMENTS ON THE STRENGTH OF A GENERALIZED INSTRUMENTAL RESPONSE

CHARLES C. SPIKER[1]
Iowa Child Welfare Research Station

The importance of the principle of stimulus generalization in behavior theory, and especially in stimulus-response theories of behavior, is generally recognized. The principle has been used to explain such diverse phenomena as the occurrence of sudden or insightful solutions of problems, and the maladaptive persistence of behavior that is not rewarded or is even punished. Despite the crucial role that stimulus generalization has played in stimulus-response accounts of behavioral phenomena, important empirical data are still missing from the literature (12).

The principle has recently taken on added significance through its application by experimenters to the explanation of results obtained from studies of the discrimination learning performance of children (1, 2, 3, 4, 7, 8, 13). In some of these experiments, the related principles of mediated generalization and acquired distinctiveness of cues (11) have also been applied. Thus, it is desirable to devise experimental conditions and apparatus that are appropriate for children and which may also be used to verify and extend some of the important findings in the area, now based primarily on the study of infrahuman and human adult subjects.

One of the important variables assumed to affect the generalization of a response is the number of reinforcements given on the training stimulus. Razran (12) reports that the evidence from classical conditioning with infrahuman and human adult subjects demonstrates fairly conclusively that an increase in the number of reinforcements on the training (conditioned) stimulus results in an increased amount of generalization to the test stimuli. A recent study with rats by Margolius (10) reports an increase in the strength of a generalized instrumental response with increased numbers of reinforcements on the conditioned stimulus.

PROBLEM

The present experiment is concerned with the development of a procedure adequate to study generalization in preschool children, and to test the explicit hypothesis that an increase in the number of reinforcements on the training stimulus results in increased strength of the generalized

---

[1] The writer is indebted to his research assistants, Mr. Lewis P. Lipsitt and Miss Kathryn Norcross, for assistance in the collection of the data.

CHILD DEVELOPMENT, Vol. 27, No. 1 (March, 1956)

response. The method whereby the conditioned stimulus is presented and the subject (*S*) is trained to make a single response was rejected when preliminary work with children proved it to be rather insensitive to variations in the amount of training. On the other hand, the operant conditioning technique, used with children by Warren and Brown (16), was also rejected since it does not allow for convenient manipulation of stimulus similarity. Instead, a combination of these two methods was developed, in which a stimulus is presented for a set period of time and, during the stimulus presentation, *S* is allowed to respond repeatedly with reinforcement following each response. Response strength is assumed to be reflected in the number of responses made during the presentation interval. This assumption appears tenable in view of the tendency for the *S*s to chain the responses: that is, after a few trials, they tend to pull repeatedly without checking for the occurrence or nonoccurrence of reinforcement until the stimulus has disappeared.

In this procedure, then, *S* learns to pull the lever repeatedly during the period that the conditioned stimulus is present and to cease pulling the lever when it disappears. Following training on this problem, he is presented with the test stimuli, and his tendencies to pull the lever to these stimuli are observed.

METHOD

Subjects

The *S*s were 36 children from the Iowa Child Welfare Research Station preschool laboratories, ranging in age from 4 years, 3 months, to 5 years, 6 months. Eighteen *S*s were assigned at random to one group, and were given 12 reinforcement trials (Group 12); the other 18 were assigned to a group given 24 reinforcement trials (Group 24). Within each of these two groups, 6 *S*s were assigned to each of three subgroups that differed only in the order in which the test stimuli were presented. Twelve additional *S*s were run in the experiment but were eliminated for reasons listed in the section describing procedures for preliminary training.

Apparatus

The experiment was conducted in a semi-darkened experimental room with an adjoining observation room. The apparatus was placed in the window between the two rooms, approximately 33 in. above the floor. *S* operated the apparatus from the experimental room while the experimenter (*E*) operated the controls and recorded *S*'s responses from the observation booth. As *S* viewed the apparatus, it consisted of a gray metal panel 20 in. wide and 12 in. high. Three in. from the top was centered an aperture, 2 in. in diameter, covered with flashed opal glass. A ¾ in. copper tube, 5 in. below and 4 in. to the left of the center of the aperture, led to a transparent container fastened at the base of the apparatus. A metal rod with a hand grip extended from the apparatus 5 in. below and 4 in. to the right of the stimulus aperture. Directly below the apparatus was a wooden stand, 2 ft. square and 8 in. high, on which *S* stood during the experiment. The stand served the double purpose of restricting the child's movements during the experiment and raising him to a level convenient for operating the lever and observing the stimuli.

The apparatus was enclosed in a metal cabinet, 12 in. high, 20 in. wide, and 12 in. deep. The front face of the cabinet was that presented to *S*. Approximately 7 in. behind the aperture were four 75-watt projection lamps, arranged so that their beams focused on the aperture. Each lamp was enclosed in a reflecting container on the front of which a 2 in. square glass color filter could be inserted. A solenoid-operated marble ejector was mounted inside the front face of the cabinet. The marble magazine, containing 200 marbles, extended from the top of the cabinet and was mounted on the walls of *E*'s control room.

The circuits to the projector lamps ran through a 10-bank, 25-contact stepping relay (Rotary Type 45, Automatic Elec. Sales Corp.), which in turn was controlled by two Hunter Decade-Type Interval Timers (6). One of the timers controlled the presentation interval and the other controlled the interval between presentations. A multi-contact selector switch was wired into the circuit in such a way that any of the 10 banks on the stepping relay could be used. These banks determined which stimuli would be presented and the order in which they would be presented.

Each pull of the response lever over a distance of ¾ in. activated one of four electrical impulse counters, provided a stimulus lamp was on at the time the response occurred. The circuit was arranged so that a response to each stimulus was recorded on a separate counter. The number of responses made to a given stimulus was readily obtained for each presentation. Approximately 8 pounds of pressure were required to move the lever the ¾ in. necessary to activate the counters and the marble ejector. The marble ejector could be activated only on those trials in which the training stimulus was present.

Stimuli

The stimuli were lights of different hue obtained by shining the beams of the projector lamps through glass color filters obtained from the Corning Glass Co. The training stimulus was a deep blue (Filter No. 5543), the first test stimulus was a light blue (No. 5031), the second was a blue-green (No. 4060), and the third was a green (No. 4010). The stimuli were not equated for brightness, Nos. 5543 and 4010 appearing to be considerably dimmer than the other two.

Preliminary Training

S was brought into the experimental room, shown a variety of dime-store toys spread out on a small table, and allowed to choose the toy that he wished to have. He was then informed that he must pay for the toy with marbles and that *E* would show him how to get some marbles. *E* then turned on the apparatus, setting it to present a series of training stimuli. This series consisted of 4-second presentations of the training stimulus separated by 2½-second inter-presentation intervals. *E* pulled the lever several times while the training light was on, pointed to the marbles that were being ejected, and instructed *S* that pulling the lever made the marbles fall into the container. Following this demonstration, *S* was allowed to pull the lever to obtain marbles. The preliminary training was continued until, for several consecutive trials, *S* began to pull more than once per light and stopped pulling during the inter-trial interval. Twelve *Ss* were eliminated from the experiment, six in each reinforcement group, either because they did not pull more than once per light, because they persisted in pulling between lights, or both. The number of subjects eliminated was due, in our best judgment, to the failure of the experimenter to repeat the instructions and demonstrations during the preliminary training. In a second experiment (15), the instructions and demonstrations were repeated until each subject met the necessary criteria and no subject was eliminated. As soon as *S* had met the criteria for the preliminary training, *E* stopped the apparatus and set the stepping relay for the starting position in preparation for the reinforcement training.

Reinforcement Training

E returned *S* to his stand before the apparatus, informed him that he was going to the next room to start the apparatus, and that *S* would be told when he had obtained enough marbles to buy his toy. *E* next returned to the observation booth and started the apparatus. The reinforcement stimulus series was like that for the preliminary training, except that *Ss* in Group 12 were given 12 presentations of the training stimulus while those in Group 24 were given 24 presentations. *E* recorded on prepared sheets the number of responses *S* made to each stimulus presentation.

Generalization Test

Following the last trial on the reinforcement training, *E* turned the selector switch so that the generalization stimulus series was presented. The time between the last reinforcement trial and the first generalization trial was nine seconds. The longer interval was necessary due to the extra contact point on the stepping relay. The generalization series was introduced with no other changes in the procedure.

The series consisted of 24 stimulus presentations. Beginning with the first, the training stimulus was strictly alternated with a test stimulus. Thus, there were four cycles, each cycle consisting of three presentations of the training stimulus and one presentation of each of the test stimuli. Three generalization series were used and they differed only in the order in which the test stimuli were presented. Considering the three series together, each test stimulus appeared equally often in each temporal position. The number of pulls was recorded for each presentation.

RESULTS

An analysis was made of the mean number of responses per trial for the 12 trials of Group 12 and the last 12 trials of Group 24 during the reinforcement training. The mean for Group 12 was 4.32 and, for Group 24, it was 4.97. The difference between these means was nonsignificant ($p > .50$). Thus, the additional 12 reinforced trials given Group 24 did not significantly increase its rate of responding over that for Group 12.

The major analysis of the data involved the number of responses on each presentation of a test stimulus. The data were analyzed by means of an extension of Lindquist's (9) Type VI analysis of variance. The four main factors were the number of reinforcements during training, the stimulus order during the generalization tests, the cycles of the generalization stimulus series, and the different test stimuli.

Table 1 gives the means for Groups 12 and 24 for each stimulus presentation. Since the order of presenting the test stimuli proved nonsignificant,

TABLE 1

MEAN NUMBER OF RESPONSES TO EACH PRESENTATION OF THE TEST STIMULI

Cycle No.	*Test Stimulus*	*Group 12*	*Group 24*	*Combined Groups*
I	1	1.61	2.33	1.92
	2	1.78	2.39	2.08
	3	1.28	2.50	1.89
	Mean	1.56	2.41	1.98
II	1	1.11	2.56	1.83
	2	1.17	2.33	1.75
	3	1.06	1.83	1.44
	Mean	1.11	2.24	1.68
III	1	0.89	2.44	1.67
	2	0.83	1.56	1.19
	3	1.22	1.39	1.30
	Mean	0.98	1.80	1.39
IV	1	0.44	1.67	1.06
	2	0.39	1.22	0.80
	3	0.39	1.44	0.92
	Mean	0.41	1.44	0.91
Total (All Cycles)	1	1.01	2.25	1.63
	2	1.04	1.87	1.46
	3	0.99	1.79	1.39
	Mean	1.01	1.97	1.49

the data in Table 1 are not broken down according to order. In Table 2 is presented an abbreviated summary of the analysis of variance. Only the main effects are presented, since none of the interactions prove significant at less than the 10 per cent level. The highly significant main effect for cycles indicates that the number of responses made to the test stimuli decreased as a function of the number of previous test stimuli presented. The significant main effect of reinforcement ($p < .01$) reflects the greater number of responses made by Group 24 as compared to those made by Group 12. The differences among the means of the number of responses to the different test stimuli proved significant only at the 20 per cent level.

TABLE 2

ABBREVIATED SUMMARY OF ANALYSIS OF VARIANCE OF RESPONSES TO TEST STIMULI

Source	*df*	*Mean Square*	*F*	*p*
Subjects	35			
Reinforcement	1	99.19	7.97	.01
Order	2	34.42	2.77	.10
error (b)	30	12.44		
Within Subjects	396			
Cycles	3	21.76	16.12	.001
Stimuli	2	2.26	1.67	.20
error (w)	330	1.35		

NOTE.—All interactions have been omitted from the table since none was significant at less than the .10 level of confidence.

DISCUSSION

The results indicate that there is a relationship between the strength of a habit, defined in terms of the number of reinforcement trials, and the strength of the generalized tendency to other stimuli. This relationship is predicted by Hull's (5) and by Spence's (14) theories of generalization, both of which are based primarily on the study of infrahuman behavior in simple conditioning situations. The relationship appears to hold also for preschool children in a relatively complex conditioning situation. This finding is even more interesting in view of the fact that the greater number of reinforcements for Group 24 did not result in a significantly greater number of responses to the training stimulus. Thus, even though responses to the training stimulus did not show the effects of the differential numbers of reinforced trials for the two groups, the differential training was reflected in the number of responses to the test stimuli.

The reduction in the number of responses to the test stimuli as a function of increasing numbers of nonreinforced test presentations is also a typical finding in generalization studies with both infrahuman and adult human *Ss*. This reduction is presumed to represent increments in inhibitory potential for the generalized S–R connections, and is understood as one of the basic processes of extinction. The very gradual reduction in the response rate for the test stimuli in the present study suggests that the procedures used here may be of considerable value in the study of the variables affecting extinction in children. The very frequently obtained records of sudden extinction were noticeably absent from the present data.

The failure of this experiment to find a generalization gradient among the test stimuli is, perhaps, not surprising considering the fact that no attempt was made to control for the brightness of the stimuli. Thus, the test stimuli differed from each other in terms of at least two physical dimensions—brightness and hue. A subsequent study (15), using essentially the same procedures but with stimuli differing only in brightness, has shown statistically significant gradients of generalization as expected on the basis of behavior theory (5, 14).

The present experiment, then, has confirmed the direct relationship between the number of reinforcements on the conditioned stimulus and the amount of generalization to stimuli similar to the conditioned stimulus. It has also suggested a fruitful method of studying some of the basic phenomena of simple learning in preschool-aged children.

SUMMARY

The present experiment is concerned with the development of a simple learning situation that is suitable for preschool children and in which some of the basic phenomena of stimulus generalization may be studied. Specifically, the hypothesis is tested that there is a direct relationship between the number of reinforcements to a conditioned stimulus and the strength of the tendency to generalize the response to similar stimuli. Thirty-six preschool children were trained to pull a lever repeatedly to a stimulus of a given hue in order to receive marbles. They were then tested on stimuli of different hues to determine the number of responses that would be made to the test stimuli without receiving reinforcement. A group given 24 reinforcements to the conditioned stimulus prior to the generalization test averaged nearly twice as many responses to the test stimuli as did a group given only 12 preliminary reinforcements. It was concluded that the expected relationship was found to hold and that the methods used show considerable promise for the study of generalization phenomena in preschool children.

REFERENCES

1. ALBERTS, E., & EHRENFREUND, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, 41, 30-38.

2. Birge, Janet. The role of verbal responses in transfer. Unpublished doctor's dissertation, Yale Univer., 1941.

3. Cantor, G. N. The effects of three types of pretraining on discrimination learning in children. *J. exp. Psychol.*, 1955, 49, 339-342.

4. Gerjuoy, I. R. Discrimination learning as a function of the similarity of the stimulus names. Unpublished doctor's dissertation, State Univer. of Iowa, 1953.

5. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

6. Hunter, T. A., & Brown, J. S. A decade-type electronic interval-timer. *Amer. J. Psychol.*, 1949, 62, 570-575.

7. Jeffrey, W. E. The effects of verbal and nonverbal responses in mediating an instrumental act. *J. exp. Psychol.*, 1953, 45, 327-333.

8. Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, 36, 471-490.

9. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton-Mifflin, 1953.

10. Margolius, G. Stimulus generalization of an instrumental response as a function of the number of reinforced trials. *J. exp. Psychol.*, 1955, 49, 105-111.

11. Miller, N. S. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, 43, 155-178.

12. Razran, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, 46, 337-365.

13. Shepard, W. O. The effects of verbal pretraining on the discrimination learning of preschool children. Unpublished doctor's dissertation, State Univer. of Iowa, 1954.

14. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

15. Spiker, C. C. The stimulus generalization gradient as a function of the intensity of stimulus lights. *Child Develpm.*, in press.

16. Warren, A. B., & Brown, R. H. Conditioned operant response phenomena in children. *J. gen. Psychol.*, 1943, 28, 181-207.

INDIVIDUAL PATTERNS OF DEVELOPMENT[1]

NANCY BAYLEY
National Institute of Mental Health[2]

One of the perennial concerns of those who are trying to understand the young human organism is with the evaluation of his present status in the light of his earlier development, and as it bears on his future. We are concerned with the cause-effect relationships between various inherent factors and environmental situations, or between early manifestations and later development. We want to be able to predict, to know *how*, and how early we can tell what a child will turn out to be like.

To do this, we set up norms of development, based on the central tendencies and variabilities of healthy children. In applying to these norms the repeated measures of children over time, we have become increasingly aware that the growth of individuals is often unstable. In a given character, such as height, or intelligence, a child may, over a period of years, shift from high to average, to low and back to average again, as compared with his age peers. The very frequency of these shifts leads us to assume that, for the most part, they are normal and healthy patterns of growth. But we should like to understand the reasons for them, in order to insure optimal development. If we can find some condition that is operating to enhance or lower a child's efficiency or to stimulate or hinder growth, then there is the possibility of doing something to control it. But the processes of growth are very complex: it is often difficult to find a plausible event that could make the difference in the developmental trend. Furthermore, such an event may seem relevant for one child, but may make little difference in other children.

In our longitudinal studies, where we have made repeated observations on the same children over time, and charted the development of each child in relation to the norms, we find that individual patterns are the rule. It is a rare child who follows the same course in all of the observed variables through all of his growth. To explain these individual patterns, we are repeatedly being forced to re-examine the nature of development, and to reconsider the relevant conditions that can determine the course of growth in an individual organism.

If we begin by considering the early stages of the embryo and fetus, it becomes evident that the organism does not start with the discrete structures and functions that we are able to observe and classify later on in child-

---

[1] Presented as presidential address to Division 7, Developmental Psychology, American Psychological Association, New York, September, 1954.

[2] National Institutes of Health, Public Health Service, Department of Health, Education and Welfare. (Formerly at the University of California)

CHILD DEVELOPMENT, Vol. 27, No. 1 (March, 1956)

hood. Starting with the relatively simple fertilized ovum, there is, with development, a progressive differentiation of both structures and functions, with a corresponding trend in each toward independence from the others. However, this independence is never complete in the living organism, and there are gradations of degree of independency that are correlated with the amount of similarity or difference between the two factors being compared. Thus the first reflex occurs in the embryo only after a given structural differentiation has been attained. In the neonate the observable behaviors are almost exclusively reflex in nature, and for much of the first year mental development is difficult to distinguish, first from reflex responses, and then from motor coordinations. This very condition forces on us the awareness that there is continuous change in both structural and functional organization in early development. It seems reasonable to assume that as long as there is growth or developmental change in an organism there will continue to be changes in organization; that is, the relative contribution of different factors to any "whole" structure or function will shift as each factor develops, matures and declines.

An adequate frame of reference for understanding change, or differentiation and growth, must include individual differences in temporal processes, as well as structural and functional processes. That is, not only do structure and function develop and become differentiated from each other, but also they do this at different rates. These differences in timing occur for different aspects of a single organism, as well as between different organisms.

Consequently, in studying individual patterns of growth we must take all of these dimensions of change into account. They all are important both in the predictions of a child's future development and in the evaluations of his current status. Only when we understand the regular processes of changing organization will we be able adequately to assess the relative importance of specific environmental variables in altering the processes of development in an individual.

THE STUDY OF PATTERNS OF DEVELOPMENT

When we have a series of measures and observations on a child over time, as he grows from infancy to adulthood, there are several ways in which we can orient the data, in order to study his patterns of development.

Normative Age Curves of Growth

The usual way to present growth, or change, over time is to compare a child, at successive ages, with values that have been obtained for other (presumably normal) children the same ages. This may be done by presenting individual curves of growth which show the child's scores in relation to some measures of central tendency and variation. Sometimes the normative frames of reference are curves representing the means and intervals of one standard deviation from the mean. Sometimes series of percentile curves are used such as the 10th, 25th, 50th, 75th and 90th percentiles.

In either of these methods of establishing norms the data are grouped on the basis of chronological age. That is, a child is compared with other children the *same age*. It is sometimes possible, however, to set up norms on groups that are homogeneous in other respects. For example, one could group together measurements of children who are at the same stages of maturity in that factor which is being measured. Thus children of the same skeletal age may be grouped together, as being equally mature physically. Such standards are useful in determining whether a child is developing normally for his own time schedule in that and related characteristics, regardless of his calendar age.

When we have longitudinal data on the same individuals over long periods of time, there is another method of expressing growth that has been found very informative. The child's own status at some stable or well-defined period can be made the standard against which his status at other times can be compared. For example, his height at closure of epiphyses in the bones of the hand, or his intelligence score at 21 years, can be taken as the point of reference. If we call this point 100 per cent, then the fraction of that figure can be computed for the relevant measures taken at all other ages. This procedure has the advantage of eliminating individual differences in size or in capacities. It thus makes more clear-cut the study of time patterns, or growth rates, within the individual. These time patterns can then be compared for a normal sample, to learn the normal variations in prococity or retardation, or lability of the factor under consideration.

Relative Scores and Increment Scores

The methods described so far are applied to total scores or measurements in individual functions or structures. Relative scores or increment scores can be applied to any of these general methods. Customarily, relative scores are expressed by comparison with the average at a given chronological age, as standard scores or quotients (e.g., IQ or DQ). They could, however, be related to skeletal age or some other standard of maturational status. Age curves of such indices are often charted to show stability or change in relative status over time. They do not indicate growth directly: thus the constant IQ when charted by age is represented as a straight horizontal line, regardless of increasing ability. A continuously dropping standard score may, for example, be consistent with steady slow growth in actual ability.

Where growth is more rapid at some ages than at others, the pattern of change is often charted as a series of increments, or rates of change. For example, annual increments in height have characteristic changes in rates of growth, with periods of rapid growth in infancy and again at adolescence.

All of these methods yield information that is useful in evaluating a child and the course of his growth, but each is limited, and each supplements the other. Some methods are more practical, and more readily put to use. Others, more difficult, may serve their purpose if they only indicate the allowances that must be made in the interpretation of data presented in the more customary way.

The Problem of Comparability of Units of Measurement

When we wish to study the over-all development of a child, the charting of growth curves presents another problem, that of setting up comparable units of measurement for the purpose of comparing different kinds of data. We would like to compare centimeters with kilograms, and, what is far more difficult, compare absolute measurements such as centimeters and kilograms with imprecise units such as increments in intellectual ability or changing patterns of skeletal ossification. Usually these latter have been derived from observations on normal samples, reduced to the amount of change to be expected from one age to another, and expressed as mental age or skeletal age and the like. Such age norms do not take into account the facts that growth is not usually equal in intensity or extent at all ages, nor are the standard deviations of the different types of test equal or alike in their growth trends. Thus, conversion of scores into "ages" tends to minimize some evident differences in rates of development, and to exaggerate others. In the field of intelligence testing this has been considered a very real problem, and various attempts have been made to set up "absolute" scales of mental units. Such absolute scales are usually derived by comparisons of overlapping distributions of scores at successive ages, taking into account both increments with age, and standard deviations. These scores are not really absolute, but they are useful approaches in this direction.

Factors and Proportions

The proportionate contributions of various substructures or part-functions to the whole may be expressed by ratios, or other statistical devices. For example, in anthropometrics we can describe varieties of body build by relating several measures to stature: the ratio, stem length/height, differentiates between the long-bodied and the long-legged; bi-iliac diameter/height is used to separate the slender from the broad-built; weight/height also differentiates between relative stockiness and linearity of build. Statistical methods, such as covariance and factor analysis, are applied to anthropometric values and to behavior data, to show the relative contributions of several sub-scores to a total measure of physique, intelligence, or personality, and the like. All of these devices can be applied to the study of age changes in individuals by securing repeated values at successive ages wherever the relevant measures are available on the same subjects, as in longitudinal studies.

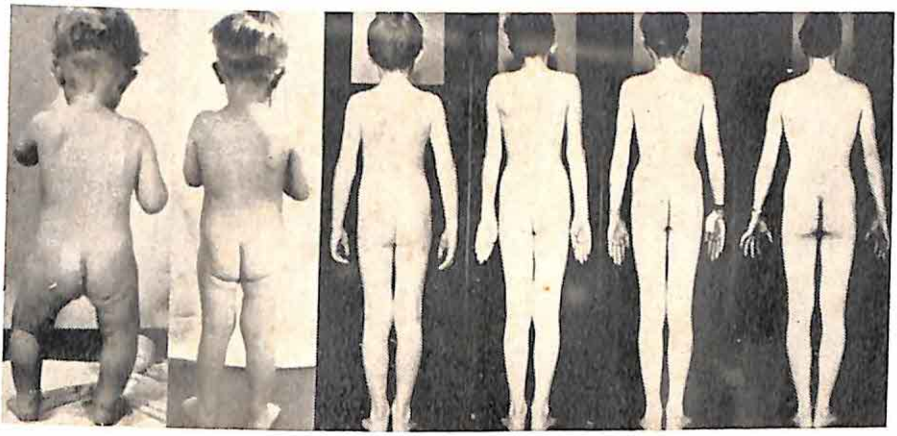


Figure 2—Changes in body proportions with growth, shown with photographs of the same boy at six ages, all adjusted to the same height.

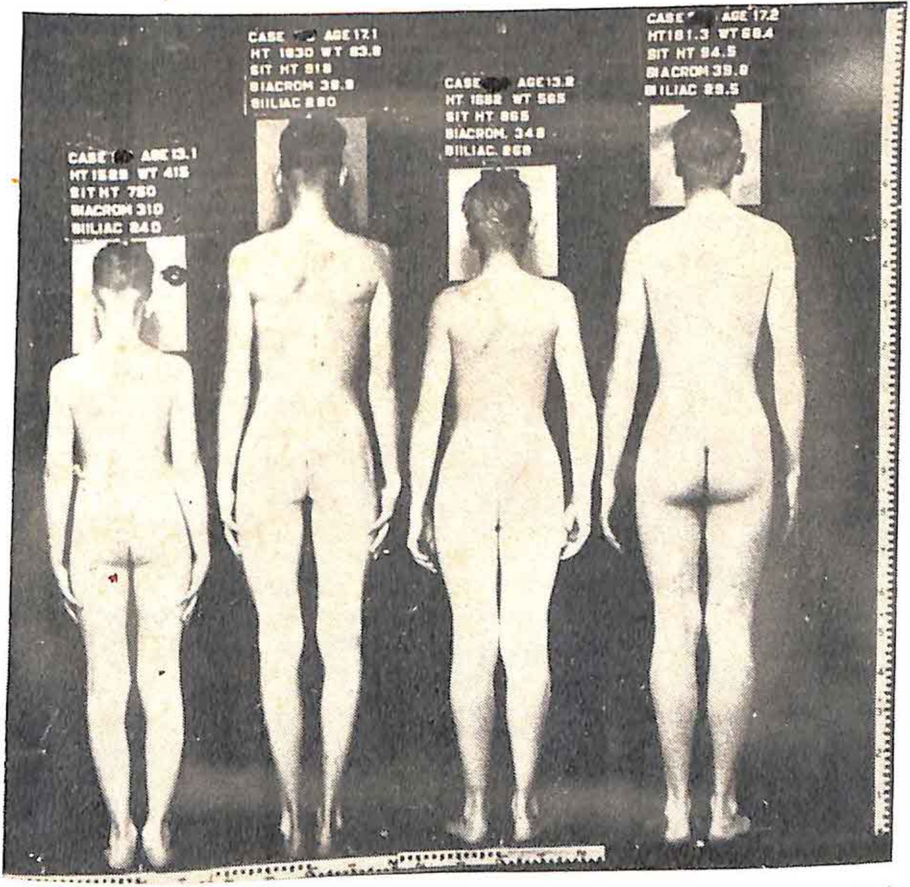


Figure 5—Comparison of two boys who differ in skeletal maturity. The retarded boy is slender-built and relatively much shorter at 13 years than is the husky accelerated boy. From Bayley (2).

DATA FROM THE BERKELEY GROWTH STUDY

I should like now to illustrate some of these methods as applied to the Berkeley Growth Study (12) data, showing their values and limitations. In the process, I hope to show, in patterns of individual development, the nature of interrelationships between several different aspects of mental and physical development from birth to young adulthood, as they are found in this small normal sample. The subjects of the Berkeley Growth Study were originally selected as normal infants when first seen in the hospital during their first few days of life. Forty of these children have continued in the study and have been tested through at least 18 and some through 25 years.

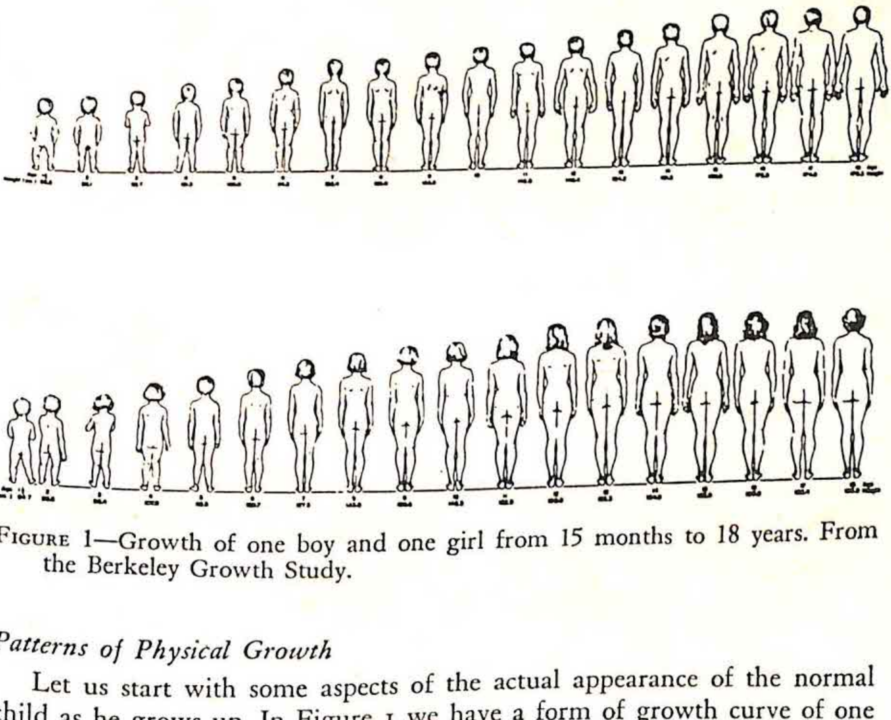


FIGURE 1—Growth of one boy and one girl from 15 months to 18 years. From the Berkeley Growth Study.

Patterns of Physical Growth

Let us start with some aspects of the actual appearance of the normal child as he grows up. In Figure 1 we have a form of growth curve of one boy and one girl from 15 months to 18 years. These are tracings from photographs, drawn to scale and spaced so that growth in size is represented at annual intervals. We see that the girl gets her height sooner, but the boy is eventually taller. Over most of this period the legs are growing faster than the stem. With maturing of the body at adolescence the girl's hips and the boy's shoulders broaden. The changes in relative proportions become more evident in Figure 2, when the photographs of this same boy, at 6 ages (15 and 30 months, 6, 11, 14 and 18 years) are reduced to a constant total height. This in effect is what we do when we compute ratios of stem length or bi-iliac width or weight to stature. The head makes up a steadily diminishing part of the total; the legs grow relatively longer to 15 years, and then the stem length increases slightly; the whole body is broad at 15 months, then slenders down, to broaden again later, mostly in the shoulder girdle. The same patterns hold for the girl, except that as she approaches maturity, it is her hips rather than her shoulders that become wide.

In Figure 3 the conventional curve of height for age is compared with a proposed new set of curves. This chart is based on data from several longitudinal studies at the Institute of Child Welfare, primarily the Berkeley Growth Study (12) and Macfarlane's Guidance Study (13). The average curve of height for chronological age of these children is represented by the heavy center line, as "smoothed" by the segments of dotted line. Actually children do not grow in height according to such means of height for age. Their growth is more like that of the solid line that changes direction more abruptly. In this chart the solid curves are based on data that are grouped, not by chronological age, but by their degree of physical maturity. Maturity was determined from readings of the X-rays of the bones by standards that are based on per cent of mature height (7). The mean age and height were determined for each successive grade of maturity. The curves on either side of the mean represent the trends of growth of children who are one *SD* or more accelerated (the curve on the left) or retarded (on the right) in their skeletal maturing. The broken lines represent the actual growth of fairly extreme cases of a tall accelerated and a short retarded child. Height curves for individual children tend actually to follow or to parallel one or the other of these lines. Figure 4 gives the same data for girls. It is notable that, except for the extremes, the mature average heights of the three groups maturing at different rates are very similar.

Body build tends to be related to rates of maturing (2). That is, the physically accelerated children are more likely to be broad-built, while the retarded children are more often slender. In Figure 5 we see typical examples of a retarded (left) and an accelerated boy at 13 and at 17 years. At 13 the husky accelerated boy is more nearly his own 17-year height than is the slender retarded boy.

In Figure 6 are shown 3 girls at maturity. They had different patterns of growth: "A" matured early, "B" was slow in attaining her final height, "C" had a very atypical rate of maturing. Their growth curves by age are shown in the left hand side of Figure 7. Slow-maturing B grew tallest, but she was relatively short before 12 years. On the right are the 3 curves of skeletal ages, as read on the Todd standards. Girl C was very accelerated from 8 to 11 years, then matured very slowly until at 13 years she fell in line with the average and maintained this status. Figure 8 gives on the left the curves for these same three girls, each in relation to herself. In attaining their own adult heights, A was advanced and B retarded, while C grew atypically. On the right are the curves of their skeletal maturity, as read on our XM standards (7) that were derived from per cent of mature height

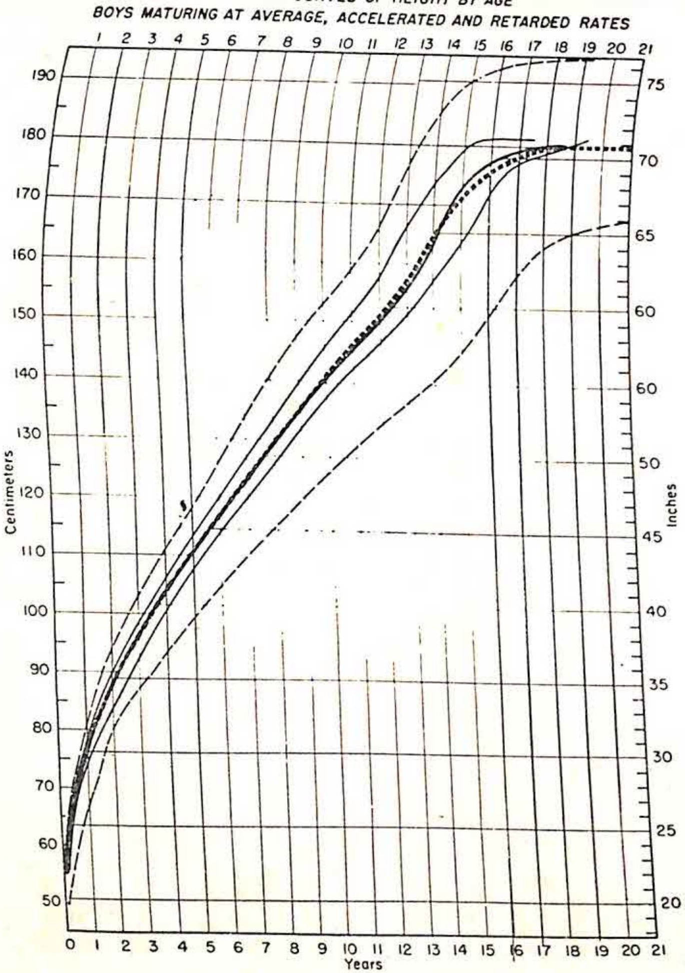


FIGURE 3—Growth curves of boys' heights, showing the difference between the conventional method of mean height for chronological age (dotted line) and height for age in groups that are equated for physical maturity.

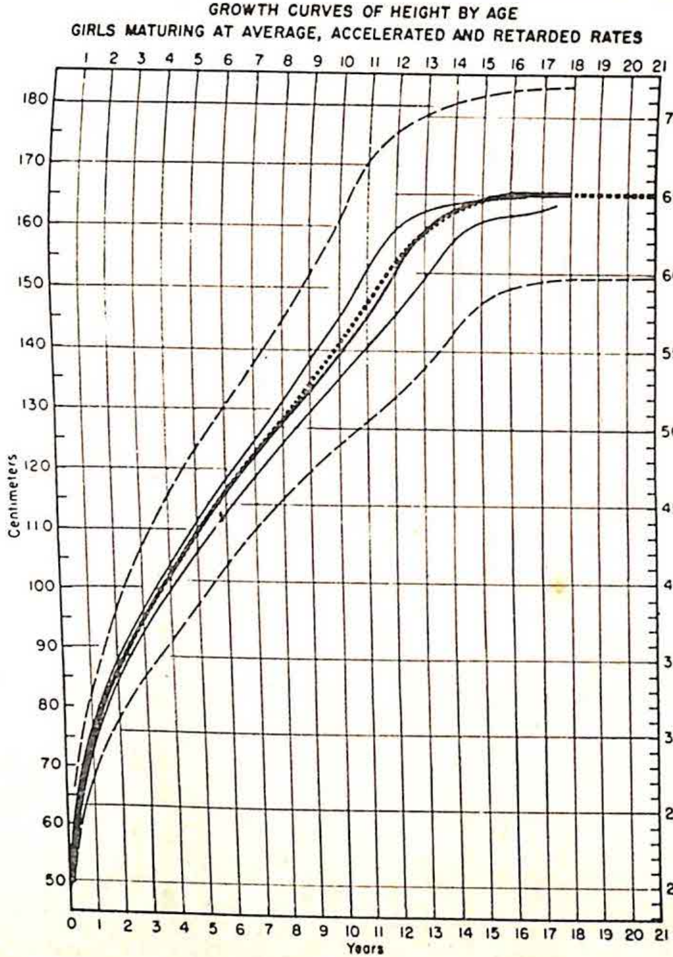


FIGURE 4—Growth curves of girls' heights, showing the difference between the conventional method of mean height for chronological age (dotted line) and height for age in groups that are equated for physical maturity

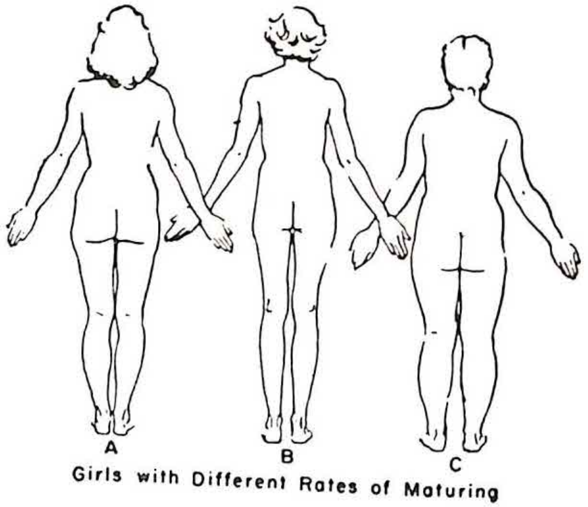

A
B
C
Girls with Different Rates of Maturing

FIGURE 6—Three girls, at 18 years, who matured differently. A was accelerated; B was retarded; C's growth was irregular.

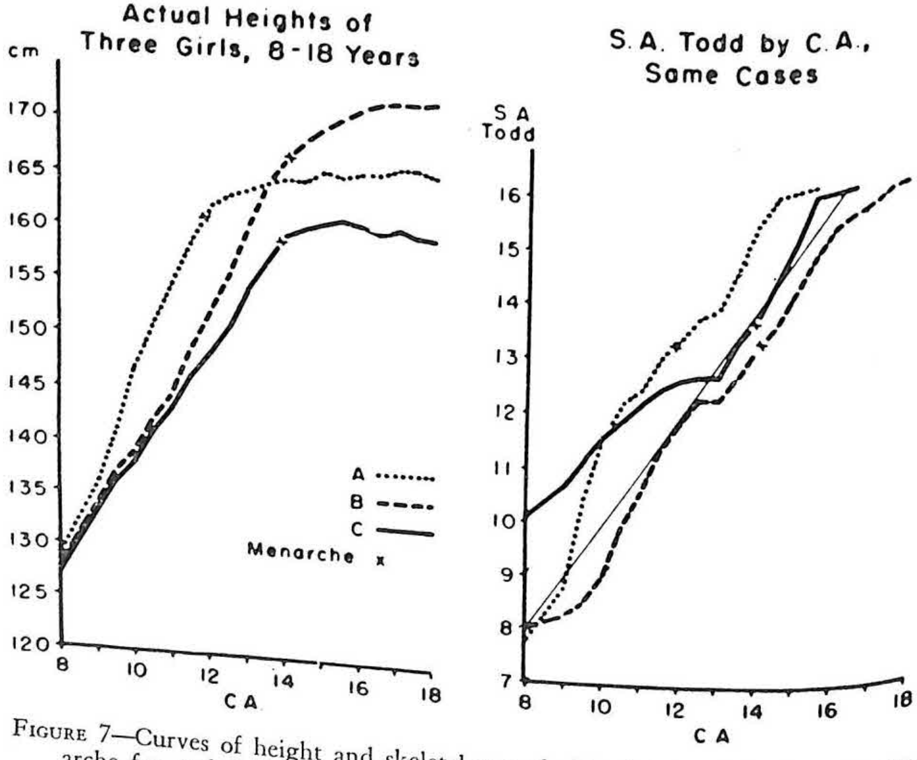


FIGURE 7—Curves of height and skeletal age of girls A, B, and C. Age at menarche for each is indicated by an *x*.

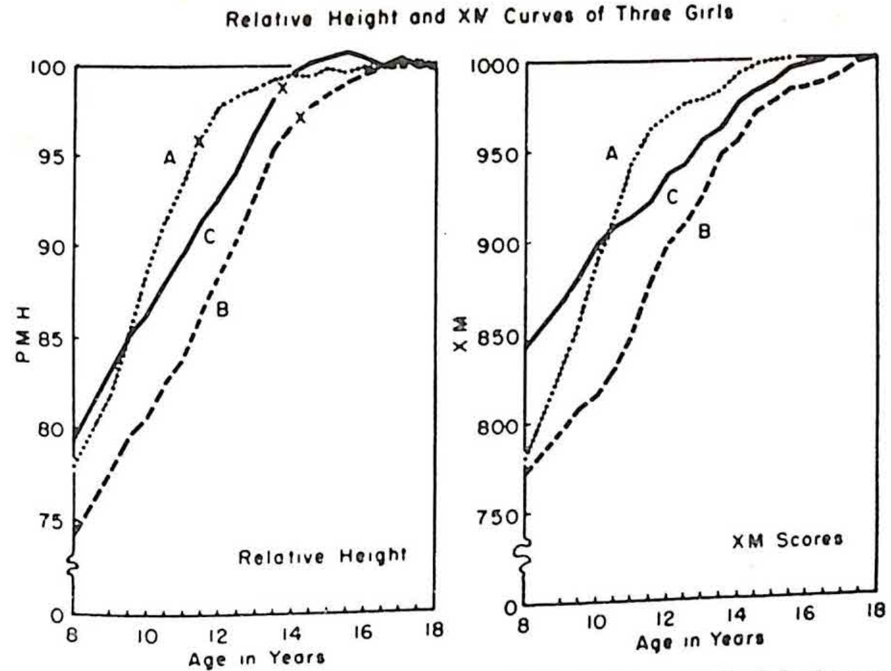


FIGURE 8—Growth curves of girls A, B, and C, showing on the left the per cent achieved by each girl of her own adult stature, for ages 8 to 18; and on the right degrees of skeletal maturity expressed in terms of per cent of mature stature.

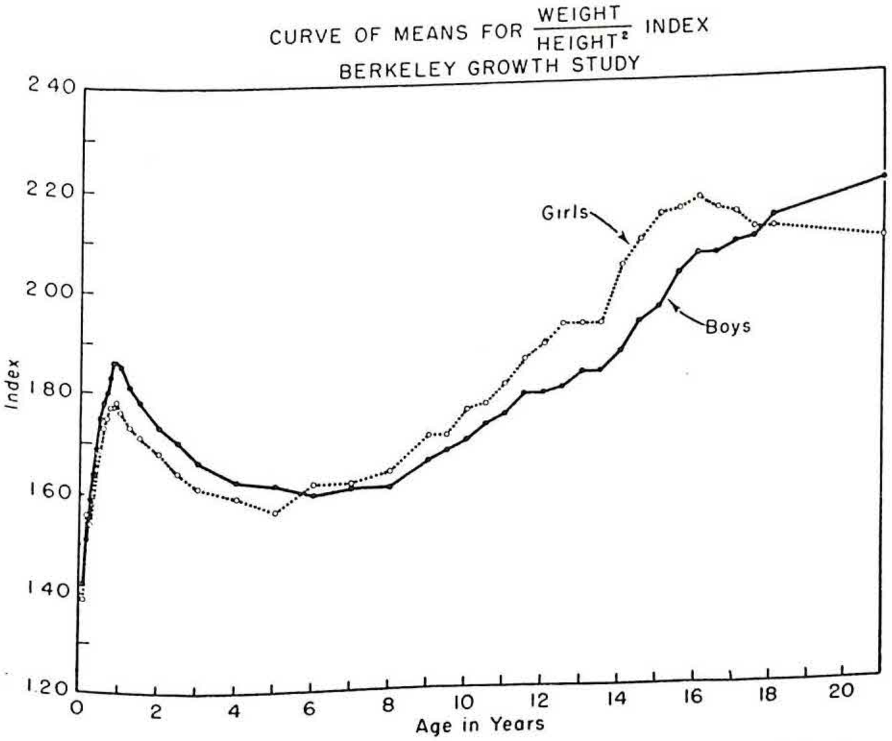


FIGURE 9—Age curves of body build, as expressed by the index W/H², for boys and for girls.

groupings. According to these curves, we see that each girl's growth is closely related to her skeletal maturing.

We have seen in the first three figures that body proportions change with growth. These changes in proportion are well represented by the age curve of the index W/H^2, as given in Figure 9 for the Berkeley Growth Study boys and girls. Proportionate to his length, the newborn infant is very slender (except for his head). He soon broadens out in all other transverse dimensions, and this trend is reflected in rapid gains in weight during the first 11 months (8). After one year a child usually slenders down to 7 or 8 years, and then broadens out again with approaching pubescence. After 16 the girls' indices drop somewhat; evidently they are losing some fat. The boys continue to gain; their "filling out" process seems to be general, including further muscular development as well as some gains in fat.

It has become evident, from these and other data on individual growth, that efforts to classify children into somatotypes, or other classes of body build, must take into account the age, or stage of their maturity. If constitutional types as exemplified in this ratio of build are stable, then we should expect the early-maturers to go through the cycles more rapidly, and the late-maturers more slowly than the average. To test this hypothesis we plotted the individual curves of the W/H^2 index for all of the Berkeley Growth Study cases. Some examples are shown in Figures 10 through 15. First is a boy who is accelerated in skeletal age. He fits the picture of a heavy-set boy who has a short period of relative slenderness. Figure 11 shows a retarded boy who is slow to gain weight. Next, in Figure 12, is an accelerated boy who was obese as an infant, and fat again later around 13 years. But we also find frequent shifts in build, such as in Figure 13. This boy was very heavy-set as an infant, but turned into a slow-maturing, tall rather slender adult. In Figure 14 is a typically heavy, accelerated girl, who contrasts startlingly with her slender, somewhat slower-maturing twin sister, as seen in Figure 15.

There is evidence from other anthropometric ratios and from series of photographs that some children's builds definitely change in some respects. Such changes seem reasonable when we consider the factors that control growth. Goldberg, an endocrinologist, has pointed out (11) that, from conception to about 2 or 3 years of age, growth is determined by the genes, which set the basic pattern. Then the anterior pituitary growth factor, in conjunction with thyroxin, becomes influential in determining growth during the childhood years. This period is essentially one of elongation of the long bones. Between the ages of 8 and 11 the pituitary gonadotropins begin to function more actively, and as androgens and estrogens appear in quantity, rapid growth is initiated. At the same time general physical maturation occurs together with sexual differentiation in physique. These three different and successively dominating controls of growth are somewhat independent of each other. They may vary both in their intensity

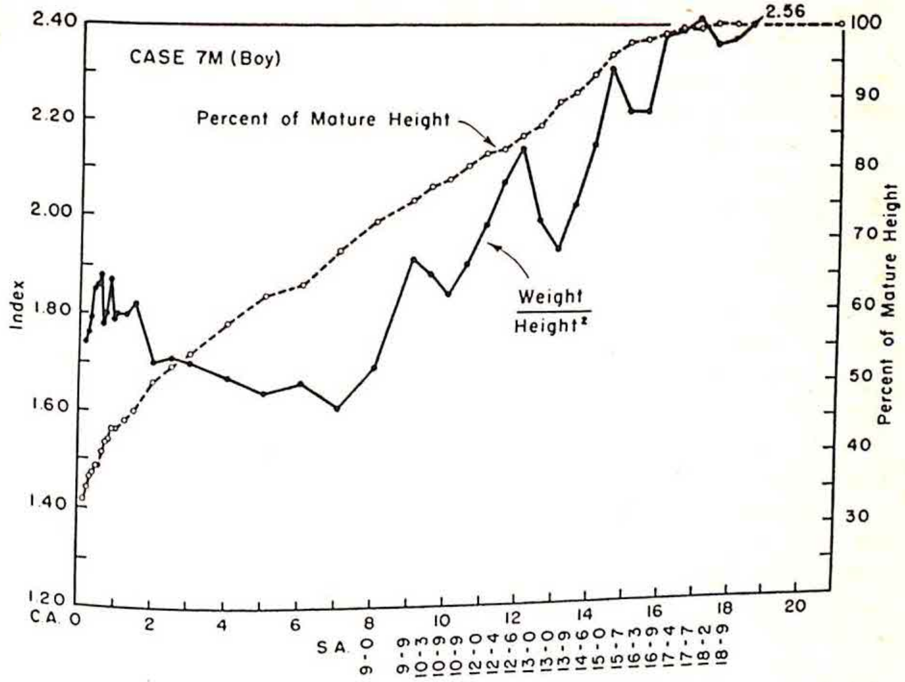


FIGURE 10—Individual age curves of body build and per cent of mature height of Case 7M, an early-maturing boy. In this and Figures 11 through 15, skeletal ages are also given below the corresponding chronological ages.

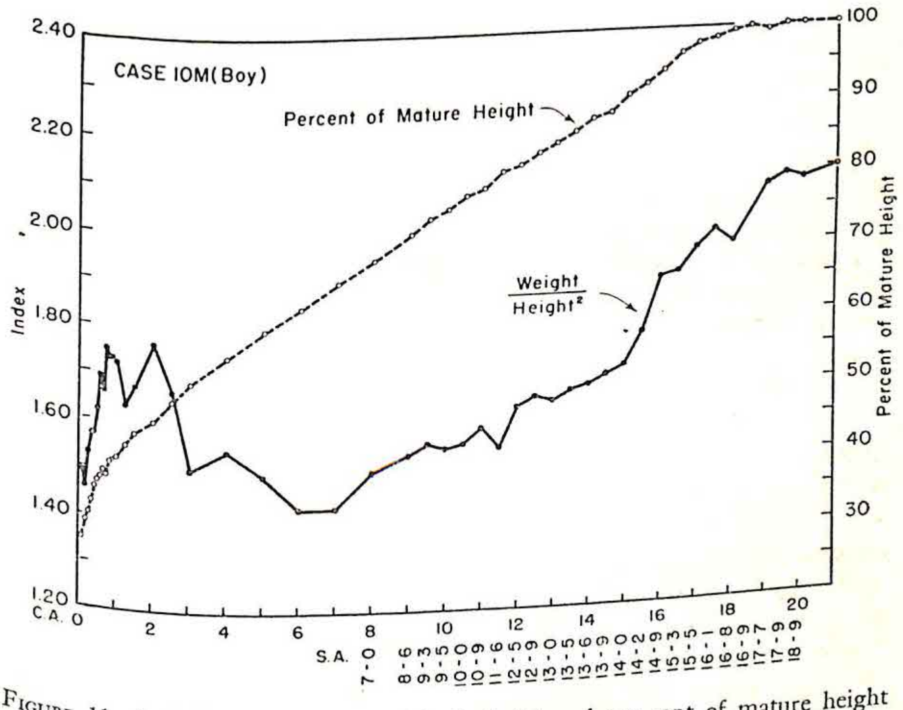


FIGURE 11—Individual age curves of body build and per cent of mature height of Case 10M, a slow-maturing boy.

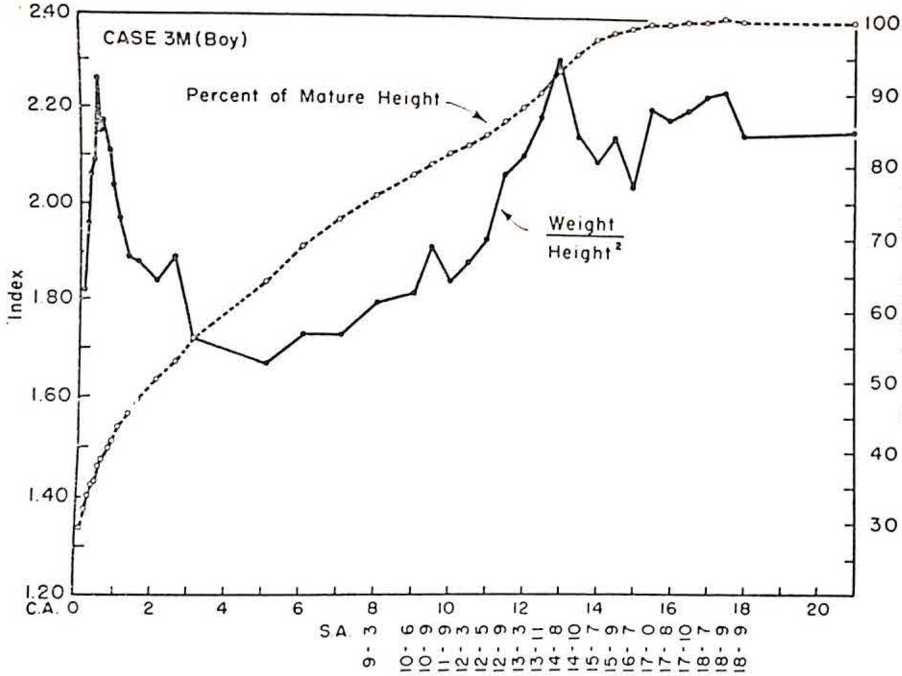


FIGURE 12—Individual age curves of a skeletally accelerated boy who had periods of exceptional weight (fat).

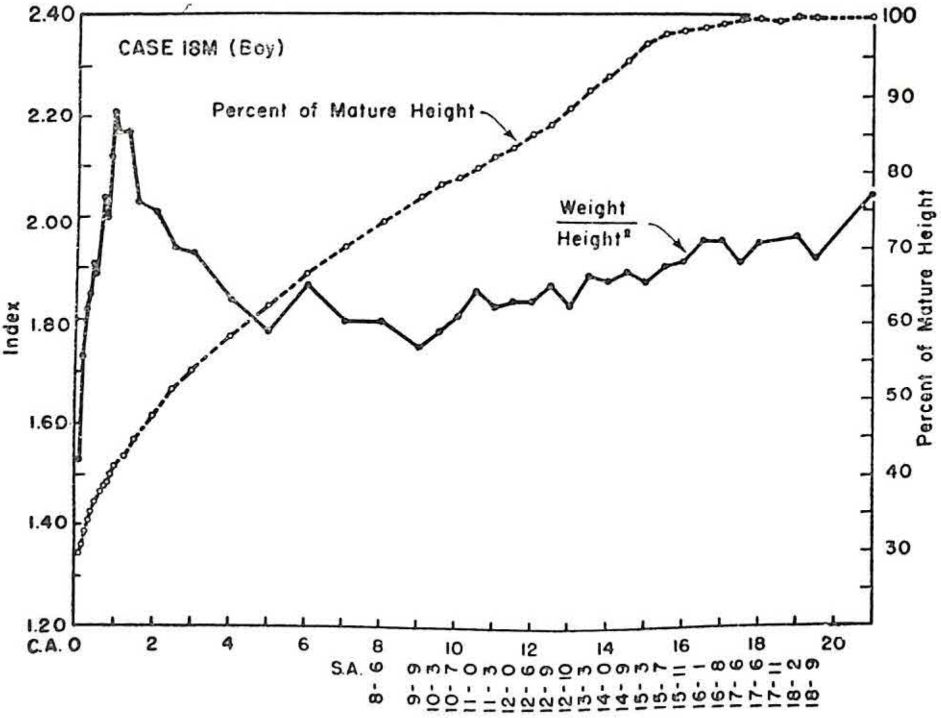


FIGURE 13—Age curves of Case 18M, a boy whose build changed from that of a heavy, stocky infant, to a tall slender adolescent and adult.

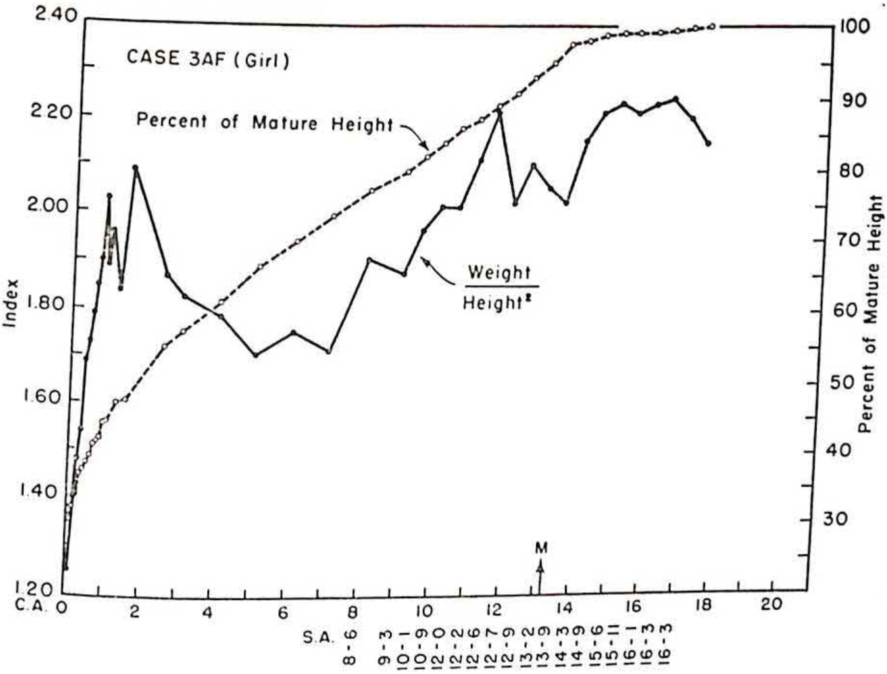


FIGURE 14—Age curves of build and per cent of mature height of an early-maturing twin (Case 3AF) who is relatively plump.

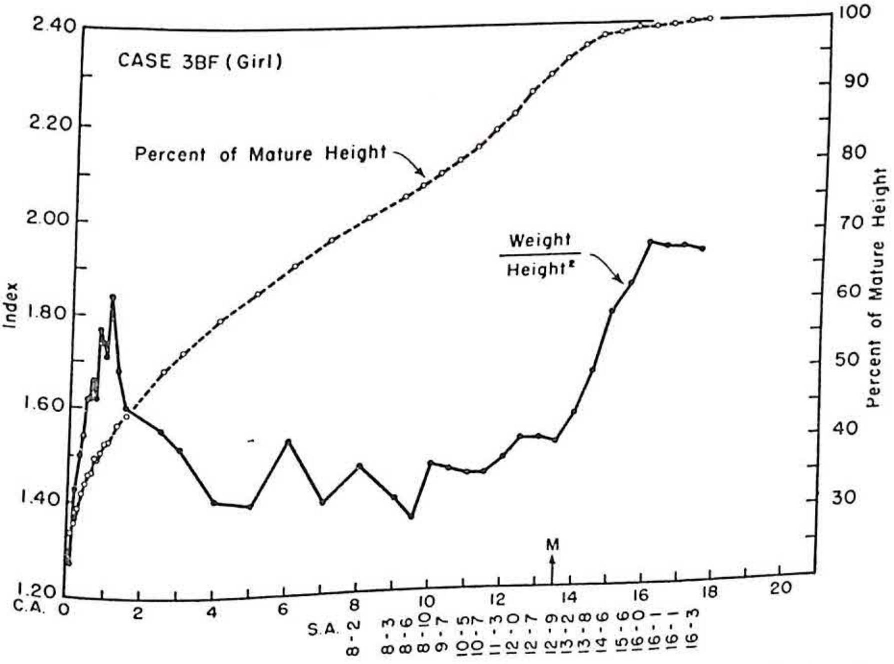


FIGURE 15—Age curves of build and per cent of mature height of 3AF's slower-maturing twin sister, who remained very slender.

and in their temporal patterns. Therefore, at the points of change in their relative influence, certain changes in characteristic growth and build may very well take place.

Age Units of Development

So far, we have been considering physical growth and maturing. Can we present the developmental aspects of behavior and intellectual functioning in ways that will be comparable, and will make possible the evaluation of mental-physical relations in development?

One method that has been used frequently is to convert all of the varibles into age units. Olson (14) has probably developed this method most thoroughly. He has assembled age norms for a variety of physical and behavioral data. From these norms, curves of mental age, height age, weight age, dental age, skeletal age, and the like, are drawn for a single child on one chart. Examples are shown in Figure 16, for years 3 to 14, of a pair of fraternal twins in the Berkeley Growth Study. These curves were prepared by Miss Lois Schulz, using Olson's norms for all pertinent data available in our records. It is interesting to compare the similarities and the differences between Twins A and B. One is more generally advanced, the other more nearly at age in most respects. Twin A is heavy-set, accel-

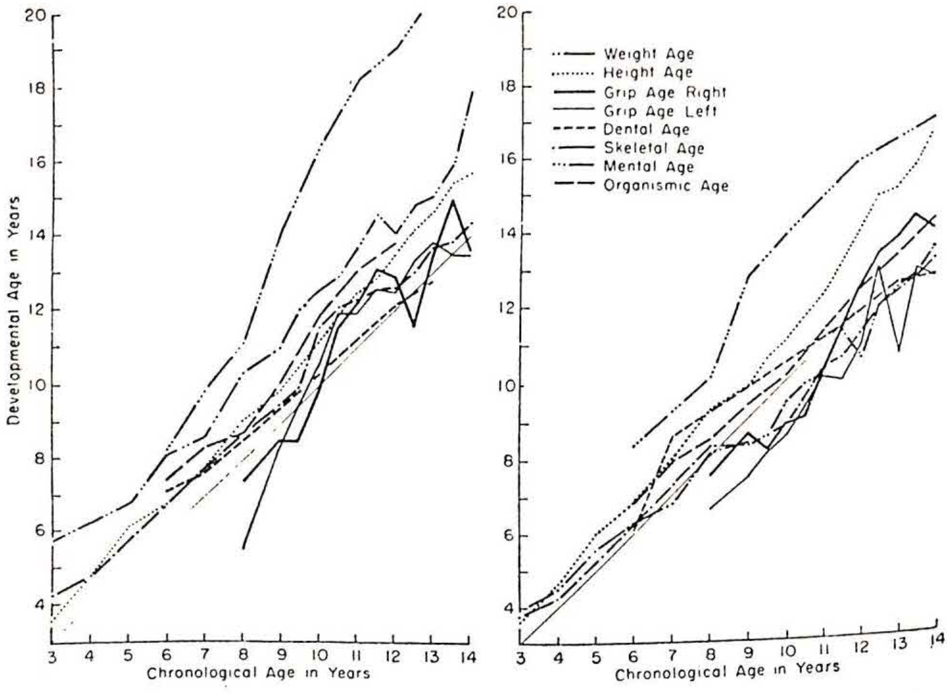

OLSON'S DEVELOPMENTAL AGES FOR FRATERNAL TWINS

FIGURE 16—Developmental age curves (Olson's method) for 8 variables of twins 3AF and 3BF.

erated in skeletal age, and markedly accelerated in mental age. After 8 years, Twin B is slender, about average in skeletal age, and generally less spread out in the different age curves. The Organismic Age that Olson has suggested is a composite of all the "ages" employed, and is intended to show whether a child is growing, in general, at an average, accelerated or retarded rate. There are, however, limitations to the usefulness of Organismic Age curves. In the first place, if one looks over a number of such charts, it becomes evident that there is great individuality in the patterns of growth, with considerable independence among the different "ages." For this reason it seems unwise to lump them together into a single Organismic Age. We may, on the other hand, learn quite a bit from studying the spread of a child's different age curves, to see where he is more likely to have trouble as a deviant, and to consider how best he can be helped to fit into a group of his age peers.

But, in the second place, we have no assurance that being advanced or retarded by a year in skeletal age means the same thing as a year's deviation in mental age. If the standard deviations of developmental age for the different variables are not similar, then the direct comparison of "ages" is misleading. Furthermore, one cannot tell whether a child is (say) advanced in weight age because he is naturally large, or because he is accelerated skeletally, or because he is obese. By some study, and comparisons of the child's differences in weight age, height age, and skeletal age, the reasons for the deviations may be deduced. But to do this requires considerable understanding of the factors involved, and they may often be overlooked.

The Quotient and the Standard Score as Measures of Relative Status

If we give up age units of growth, and turn to indices of relative status for successive ages, there are two methods that have been used: the ratio or quotient of a given developmental "age" to chronological age (of which the IQ is the best known and most used) and the standard score (or sigma score) based on the means and *SD*s of the normative sample at successive ages. The use of the ratio or quotient has the same limitations as the developmental ages on which it is based. That is, unless all of the age variables under consideration have the same standard deviation the quotients will not be comparable. Furthermore, all of the *SD*s must increase in the same constant way in order to maintain constant quotients for all ages. This is, however, not the case (4). Both mental and physical measurements, while showing increasing variability with age, generally, have periods in which the *SD*s are expanded and then contracted again.

It would seem that the way out of the dilemma is to use standard scores. In most respects, standard scores avoid the faults of the age quotient. In Figure 17, for example, are curves showing the relative mental scores of one girl from one month to 17 years (4). The solid line gives her standard scores for the Berkeley Growth Study norms; the broken line gives her IQs.

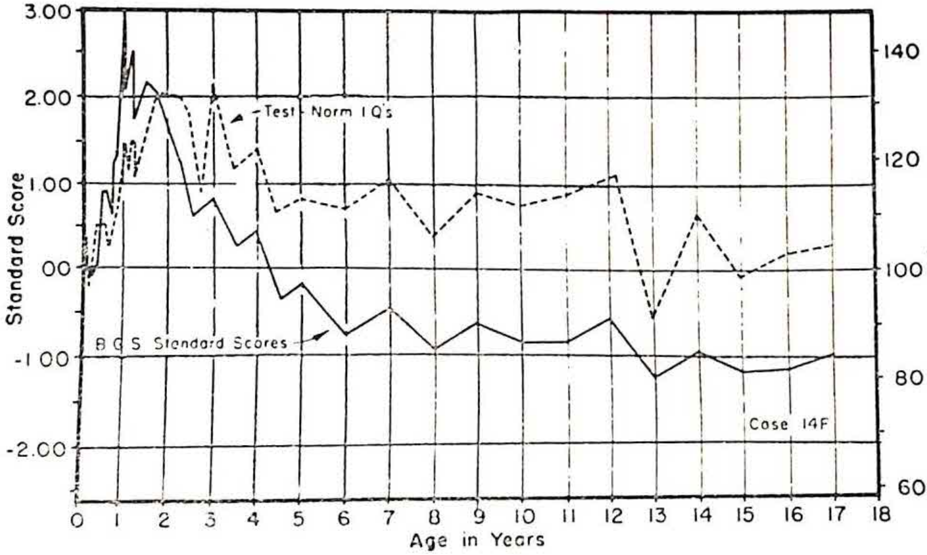


FIGURE 17—Intelligence score curves of Case 14F, expressed as IQs (broken line) and as standard scores according to the Berkeley Growth Study norms. From Bayley (4).

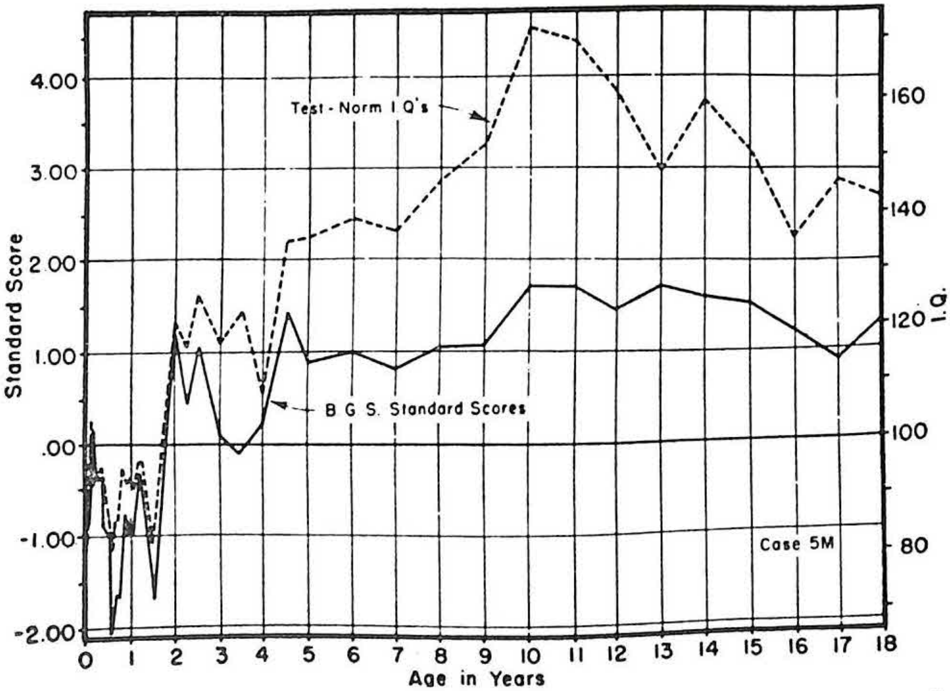


FIGURE 18—Intelligence score curves of Case 5M, expressed as IQs and as standard scores. From Bayley (4).

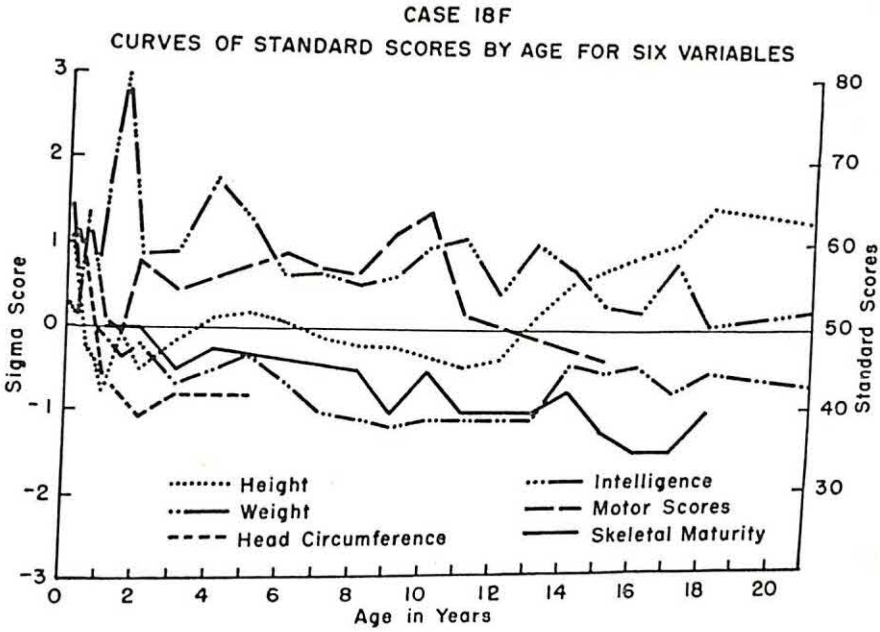


FIGURE 19—Curves of standard scores by age for six variables, for Case 18F.

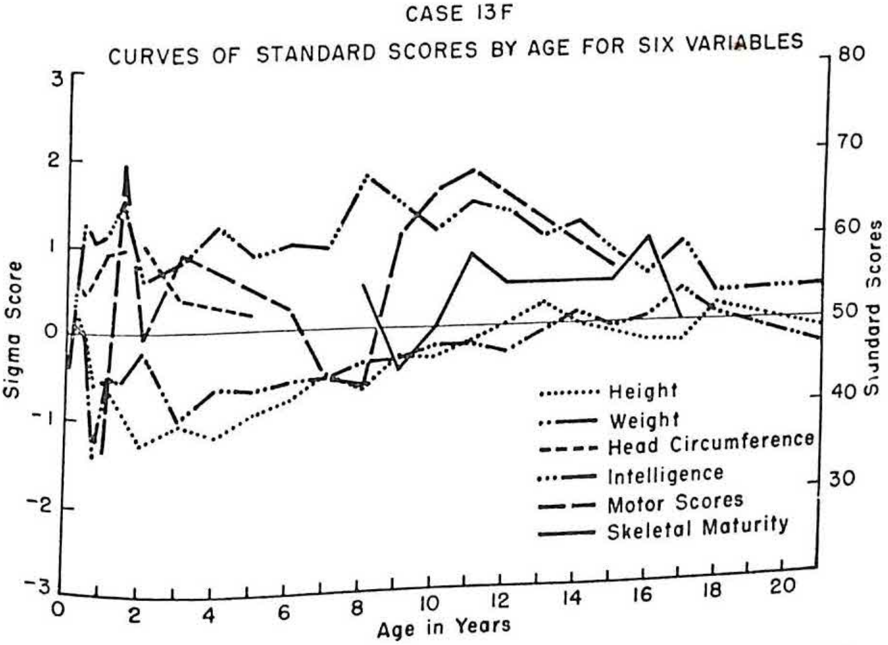


FIGURE 20—Curves of standard scores by age for six variables, for Case 13F.

At 12 months she was the brightest child in the study, and her standard scores at this period are higher than her IQs. This is because the variability in scores in the group was very small around one year. Figure 18 gives the IQ and standard score curves for a boy, whose IQ rose to a very high point at 10 years (4). His standard scores after 4 years are far more stable than his IQs.

From our growth records, we have computed standard scores for several different behavioral and physical measures. Among these are height, weight, head circumference, skeletal maturity, mental and motor development. Standard score patterns for six variables on one girl are shown in Figure 19. A detailed reading of the chart shows that this girl was small and slender, with a small head; she matured slowly and became a tall slender adult; she was consistently superior in intelligence, but more so at some ages than at others; her scores in motor tests varied between superior and average. Around one year of age, when her intelligence score was highest,

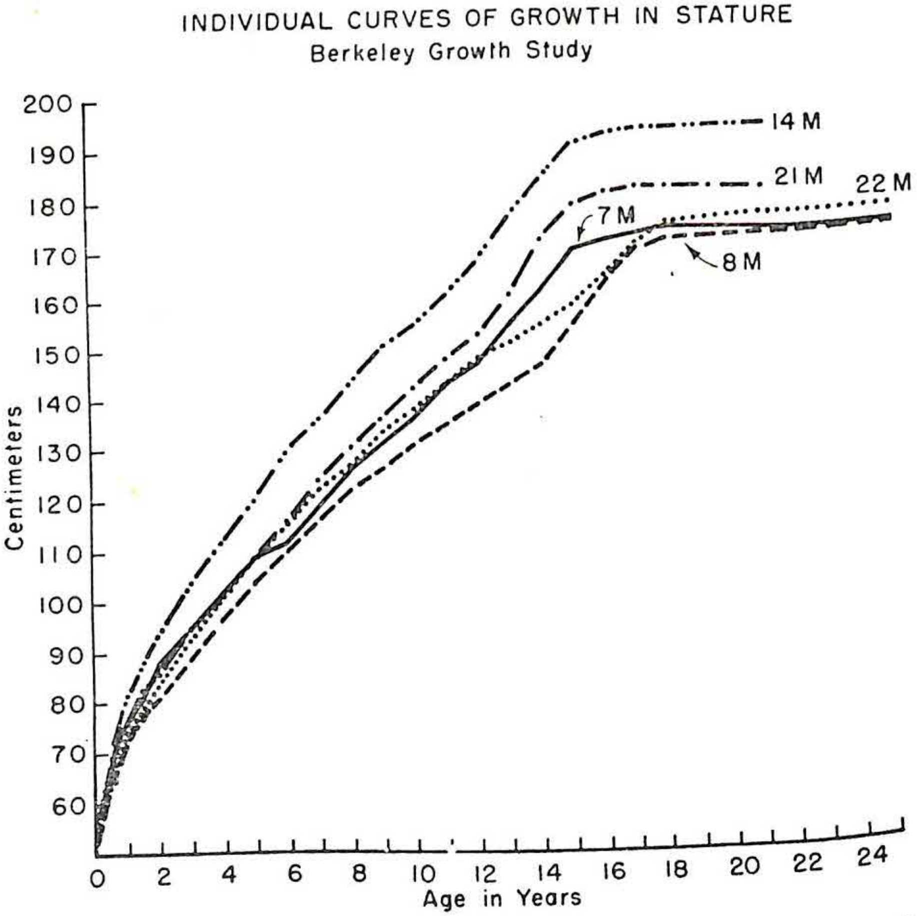


FIGURE 21—Curves of stature by age for five boys in the Berkeley Growth Study.

her pattern of motor scores appears to be almost exactly reversed. In Figure 20, case 13F presents a different pattern, with a tendency for all scores to converge toward the average as she approaches maturity. If space permitted, other cases, with equally diversified curves, could be shown.

Again, as was seen in the Olson method of Developmental Ages, these individual cases display great independence in the patterns of growth of the different variables. In some instances there appears to be concomitant variation between those variables that are highly intercorrelated, such as height and weight. Also, there is evidence that certain ones, such as mental and motor development, become more independent of each other with age.

There is one serious drawback in the use of standard scores for curves of this kind, as applied to the first year or two of life. As you may have noticed from the foregoing illustrations, the variation in standard scores during the early years is exaggerated. Because the spread of scores at these ages is actually very small as compared with the later ages, a difference of a few centimeters, or of one or two items in the mental test, can mean a

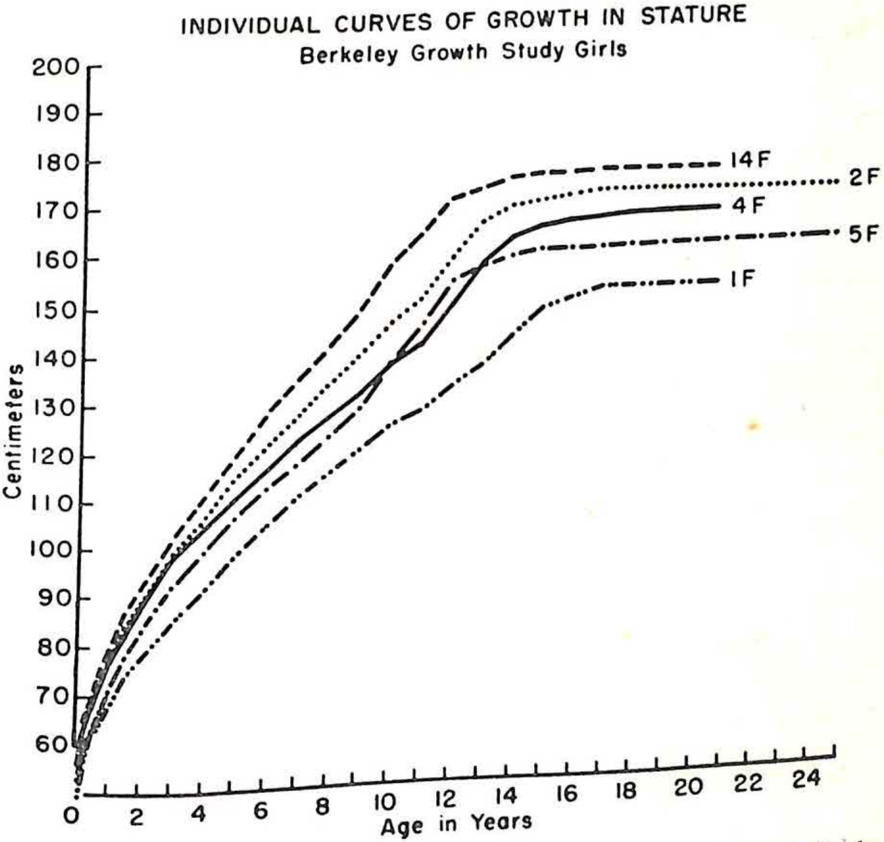


FIGURE 22—Curves of stature by age for five girls in the Berkeley Growth Study.

difference of half a standard deviation or more. In comparison with the early rapid growth, individual differences at a given age are small, and the infant's growth is much more stable than these wide swings in standard scores would imply.

Absolute Units of Growth

The relative importance of absolute units in evaluating growth and change may be seen in the next series of illustrations. Figure 21 shows the curves for height of five boys, who were selected from the Berkeley Growth Study as the tallest and shortest, at maturity, together with three of intermediate height. The heights of all five are very similar at first, as compared with later ages. Note, also, that there is some crossing of the curves later, because of differences in rates of maturing. Each individual curve is very stable, following a definite pattern. But differences in rates of maturing can make a child's standard score for height shift up and down considerably

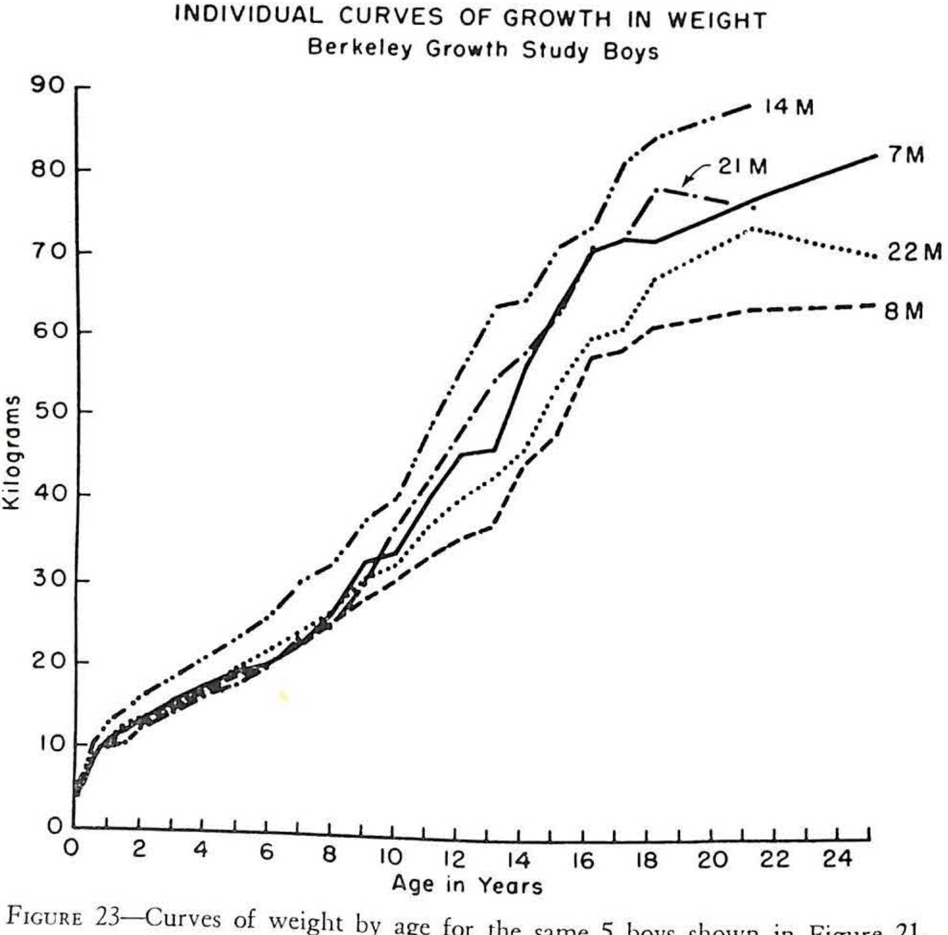


FIGURE 23—Curves of weight by age for the same 5 boys shown in Figure 21.

during the period of puberal growth. Figure 22 shows similar data for five girls.

Changing to another variable, Figure 23 presents the curves for the same 5 boys of weight for age. Again we see that at first the differences in weights are very small. The curves for weight are much less regular than for height: Weight is far more subject to short-term influences, of variations in health, nutrition and the like, than is height. In Figure 24, giving weights for the 5 girls, this instability is even more evident, and notably in the larger, heavier girls. Some of them become quite fat, and then reduce.

What about intelligence? Can we show whether or not it presents a similar trend toward increasing diversity with growth? Here we are confronted with the problem of setting up comparable units of increment for intelligence. There is ample evidence that mental ages in normal samples tend to become more variable with age. In working with the Berkeley Growth Study scores on intelligence tests the problem was further complicated by the need to compare scores from several different tests, having different methods of scoring. This has been done by the device of relating all scores to the mean and *SD* earned by this group on the Wechsler-Bellevue

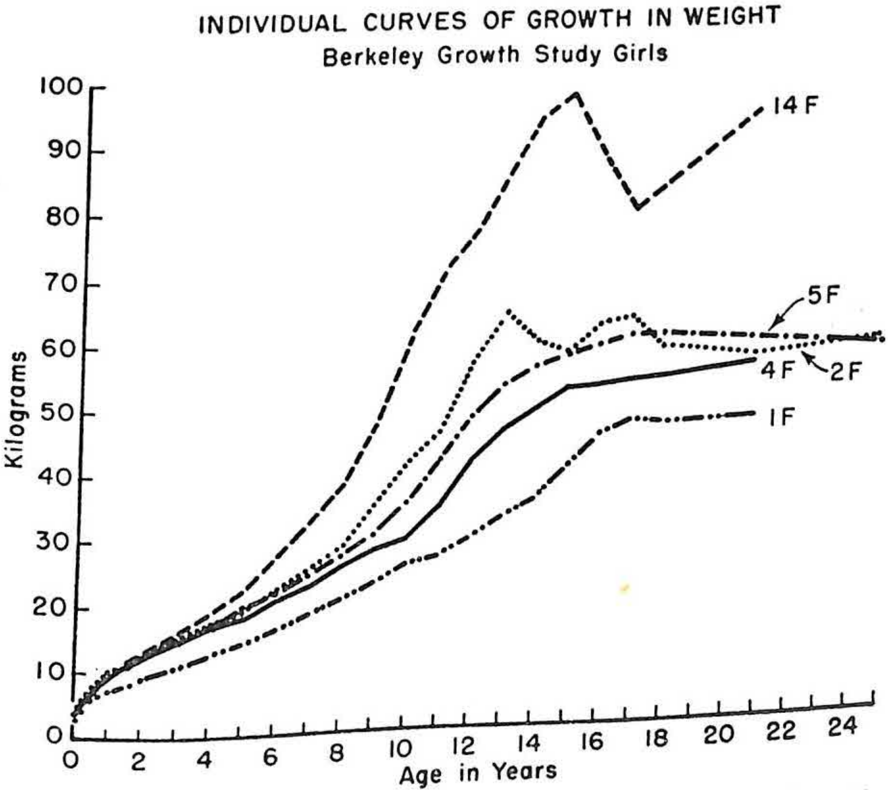


Figure 24—Curves of weight by age for the same 5 girls shown in Figure 22.

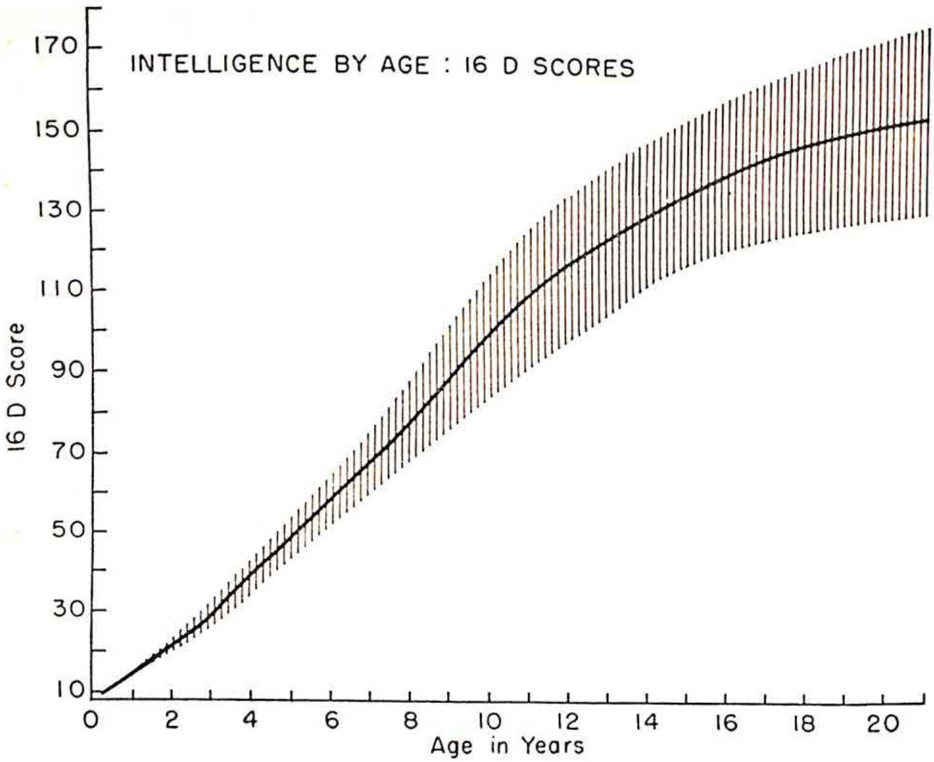


FIGURE 25—Curve showing the mean and *SD* of intelligence scores of the Berkeley Growth Study children from birth to 21 years. Scores from several different tests were converted into units based on the mean and *SD* of the 16-year test scores. From Bayley (6).

scale at 16 years (6). By adjusting the means and *SD*s earned on the Stanford-Binet and the Terman-McNemar tests at adjacent ages, all scores for the three scales and for all ages, were converted into units of deviation from the 16-year mean, in terms of the 16-year *SD*. This has been called tentatively, the 16-D score. Figure 25 gives the means and *SD*s of these 16-D scores from one month to 21 years. They are not "absolute" units, but they do give a general picture of growth relative to the status of this group at 16 years. Again the variability in scores for the first 2 years is very restricted, as compared with the later ages.

In Figure 26 are individual curves for 5 boys selected to represent wide differences in adult intelligence. Note that all five have been tested through 25 years, and in all of them the scores are still increasing. These curves, too, are less regular than the height curves, but perhaps no less regular than the weight curves. One gets the impression both of differences in rates of maturing and of differences in inherent capacity. Superimposed on these general trends are the fluctuations that may be related to differences in motivation or drive, or to other factors such as emotional distractions, or

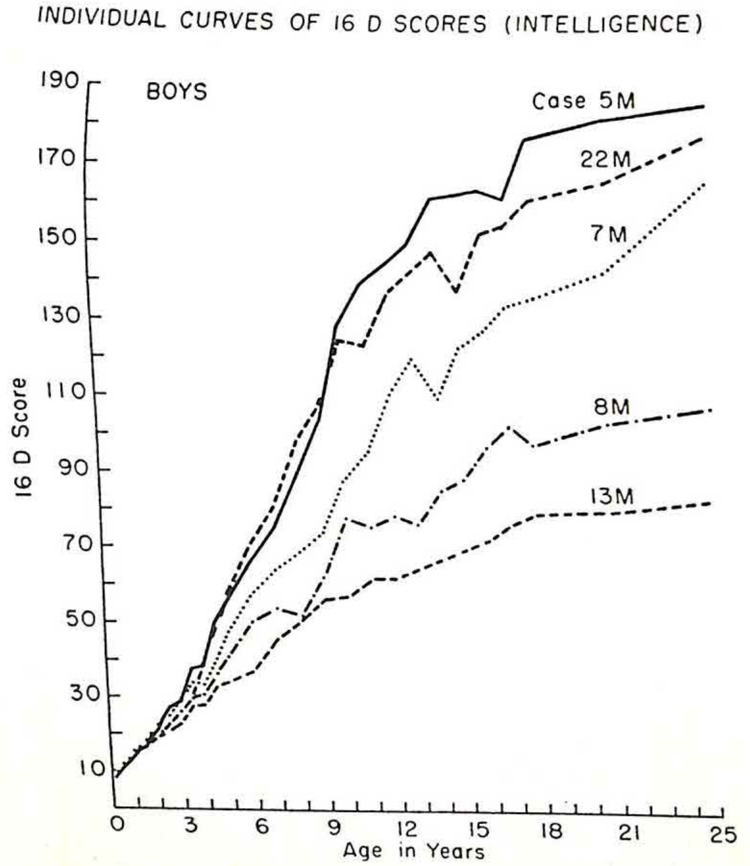


FIGURE 26—Individual curves of intelligence of 5 boys, from 1 month to 25 years. From Bayley (6).

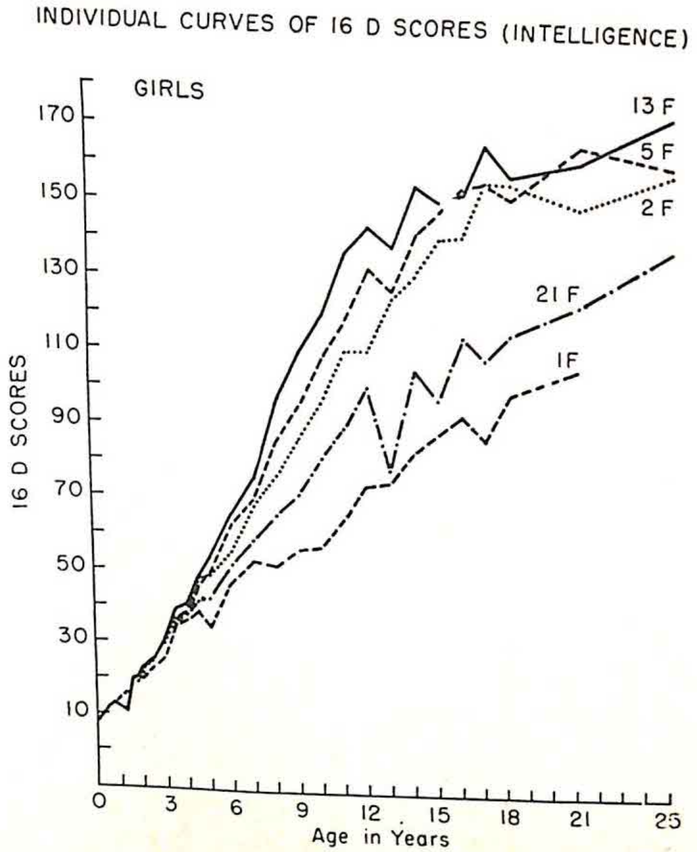


FIGURE 27—Individual curves of intelligence of 5 girls. From Bayley (6).

differences in the content of the tests used at different ages. Figure 27 shows the same thing for 5 girls. When we see how closely alike the scores are in the first two years, it seems obvious that this is at least one reason why it is impossible to predict later intelligence from scores in infancy.

Individual Differences in Rates of Maturing

It must be remembered that some of the changes in patterns of standard scores are due to the fact that children differ in their time schedule in going through the processes of growing up. In Figure 28, for example, are the standard scores for height of three children (5). They are compared with the standard scores of their own parents' height, to show that the children differ in the age at which they come to approximate the parental status. One important reason for the early deviations in growth pattern is that the children are maturing at different rates. The first, 18F, is a slow maturing girl who became tall, like her parents, only after 16 years. The second is a small boy who was like his parents from an early age. The third had very rapid growth, and was accelerated skeletally during childhood. He was tall for his age between 2 and 16 years, when he matured and stopped growing, becoming a just average-height adult. In Figure 29 we see similar variations in patterns of approaching parental status (5). This time the children's intelligence is compared with their parents' education. Possibly these patterns, too, can be explained as inherent differences in rates of intellectual maturing.

Is there some way to control or rule out individual differences in size or ability, and thus simplify the study of differences in rates of growth? As I mentioned earlier, when we have available a series of measures taken during growth, together with a stable measurement at maturity, then that stable measure can be used as a point of reference, against which any child's other measures can be evaluated. For example, we can take each child's height at the time the epiphyses of his long bones fuse and he has practically stopped growing, and call this figure 100 per cent (10). The portion of this stature attained at each earlier measuring can then be computed. In this way all of the children's rates of growth toward their final statures can be compared without the confusing differences in size. Individual curves of per cent of adult height are given in Figure 10 through 15, where they may be related to both the child's chronological age and his skeletal age.

We may try to approximate this condition for the intelligence test scores, but there are several differences between the two. As noted already, the units of intelligence are not directly comparable, and can be considered only as approximations of more and less ability. But there is another difficulty. These children did not conveniently stop improving in their intelligence test scores. Fourteen out of 15 so far tested at 25 years are still gaining. For present purposes, however, it was necessary to choose a point for comparison. Twenty-one years seemed most reasonable, as the age at which recent

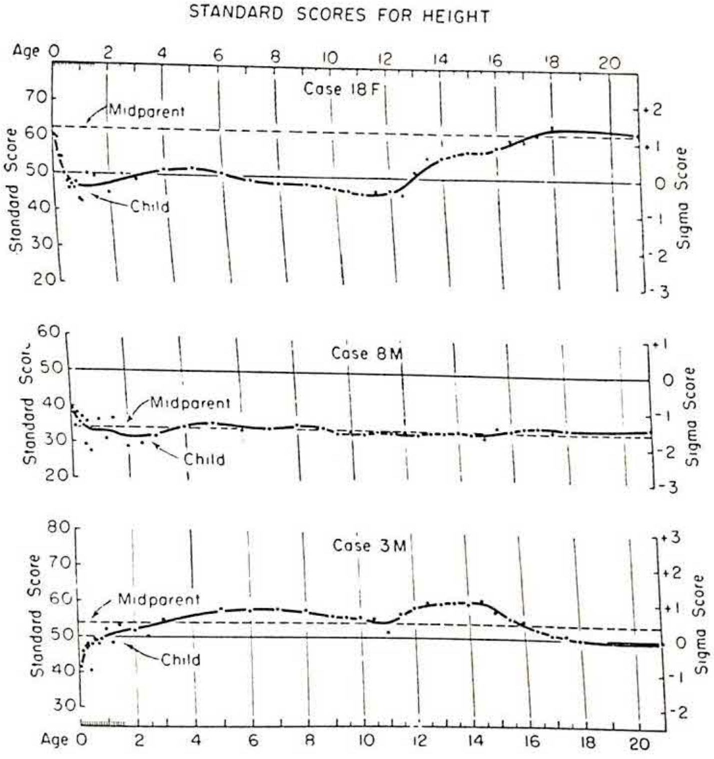


FIGURE 28—Relative heights, by age, of three children, shown in relation to the average relative height of their own parents. From Bayley (5).

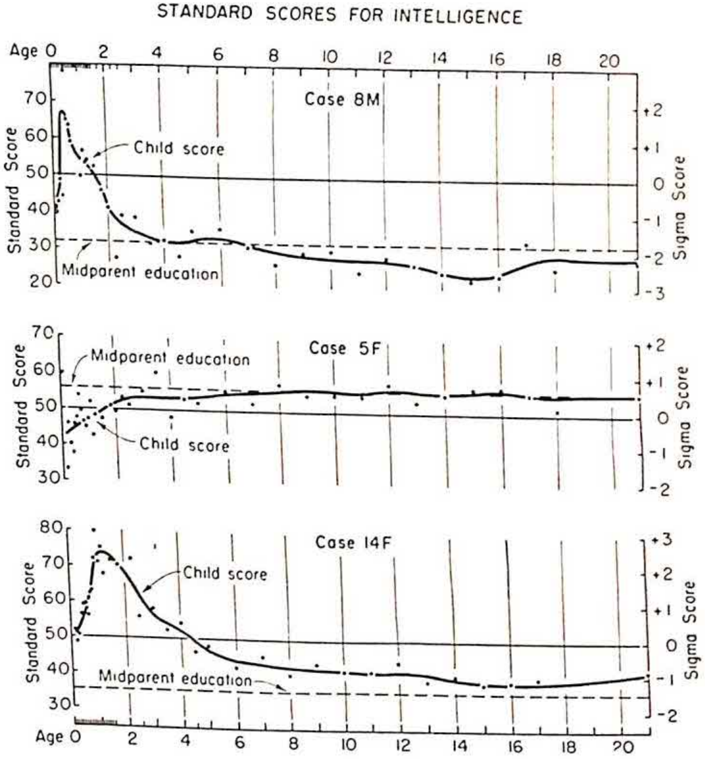


FIGURE 29—Relative intelligence scores, by age, of three children, shown in relation to the relative education of their own parents. From Bayley (5).

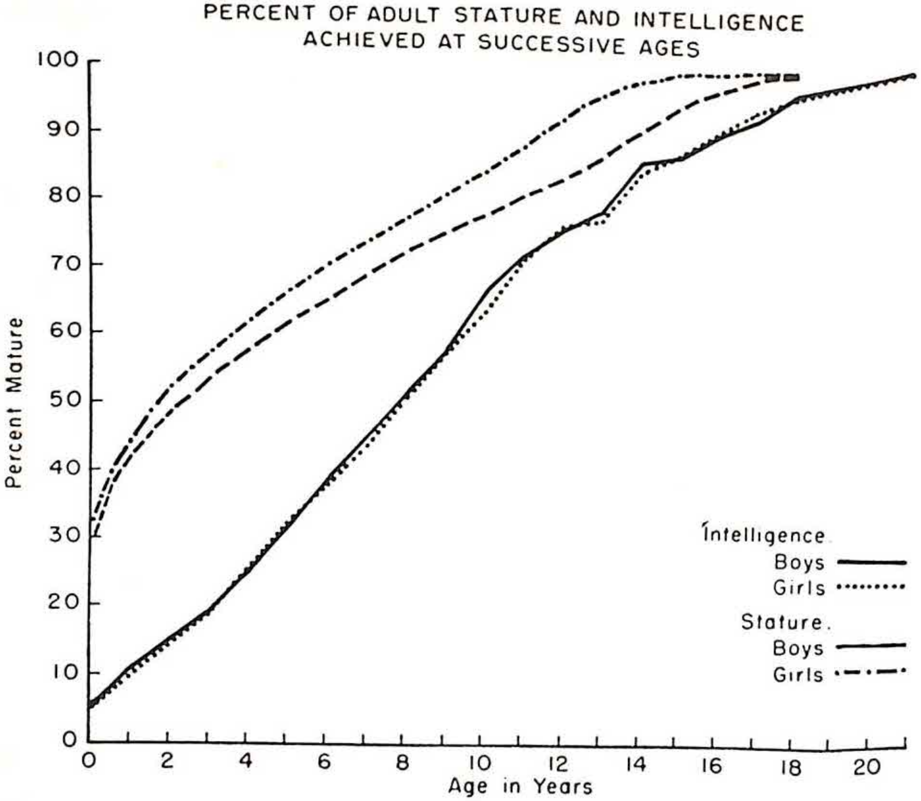


FIGURE 30—A comparison of growth in stature and intelligence, when both are expressed as fractions of mature status.

investigations have usually found the high point in intelligence test scores. Using 21 years, then, as 100 per cent, and the 16-D scores, each child's earlier per cent of adult intelligence was computed. Figure 30 gives the curves of the means for the per cents of adult status for height and for intelligence, and for boys and girls separately of the Berkeley Growth Study Sample. It is evident that growth in the two factors differs in several ways. We find obvious sex differences in one and not in the other. The periods of rapid and slow growth do not coincide, but seem almost to alternate. A study of annual increments indicates that intelligence is accelerating most rapidly up to 5 or 6 years, with another short spurt between 9 and 11. Height, on the other hand, though most rapid during the first 2 years, grows at a declining rate until 8 or 10, with adolescent spurts between 10 and 12 for the girls and 11 and 15 for the boys.

With such independence in the rates of growth in these variables we should not expect to find very closely related mental and physical maturing in the same child. A few correlation coefficients bear on this point. These children's heights and 16-D intelligence scores are positively correlated, as

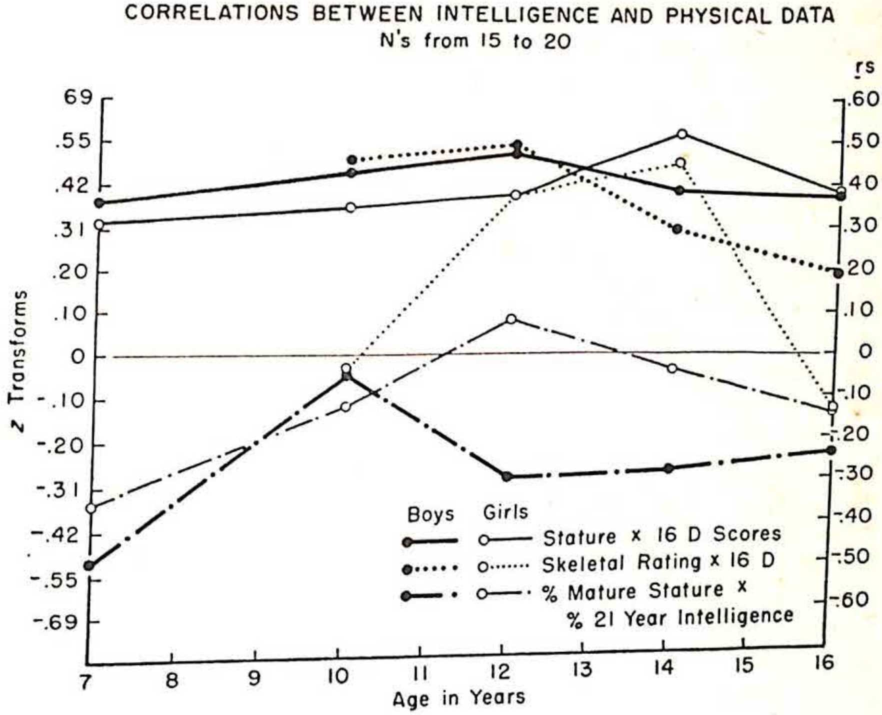


FIGURE 31—Correlations for boys and girls separately, of intelligence scores with stature and skeletal age, and of per cent of 21-year intelligence score with per cent of mature stature.

seen by the *r*s in Figure 31, connected by the solid lines. Actually, the *r*s, which vary around .40, are surprisingly high, when compared with similar correlations reported for other studies. Shown here also, connected by dotted lines, are the correlations between skeletal maturation and the intelligence scores. These *r*s tend to be positive, also, until skeletal maturity is reached.

When, however, we compare the per cent of mature height with per cent of 21-year intelligence scores, as seen in the two lower curves, the picture is very different. These *r*s are not large, but they are, with only one exception, of negative sign. In this small group of children, it seems that the more intelligent children tend to be both larger than average, and accelerated in physical maturing. But within the individual child his own mental and physical rates of growth are not concomitant. If anything, there is a suggestion that those who are slower in physical maturing approach their 21-year intelligence sooner. That is, if we think of it in the more usual way, there may be a tendency for the less able (and these are also, generally the socio-economically less favored children) to be slow in maturing physically. But they attain most of their intellectual growth early, and slow down in this respect as they approach adulthood. These same children, as I have reported earlier (1), are the ones who in infancy tended to make high scores on the mental tests. That is, the children's mental test scores for the first 6 months correlate negatively with their scores at 3 years or later. There is also a negative correlation between the scores at ages 3 to 7 months with the amount of their parents' education (9).

In this presentation, I have neglected the environmental factors, as they bear on individual patterns in development. Other areas not dealt with are emotions and various aspects of personality, which cannot easily be fitted into development sequences or scales. These factors have been omitted, in part, because it seems to me necessary first to understand the conditions under which they must be evaluated if we are to estimate their effectiveness.

It is evident that the whole process has become very complex. But, keeping in mind as many as possible of the relevant factors just reviewed, we can go down the list of cases, and try to decide which array of factors is pertinent to the course of each child's physical and mental development. Let me list briefly a few examples that will illustrate a few kinds of relationships we have found.

In a pair of fraternal twins, who were born just a little prematurely, there were closely similar patterns of rapid early growth in both size and behavioral functions. At this very early stage, structure and function were not yet independent. However, later, their mental and physical curves became very dissimilar: One twin became heavy, the other remained slender. One twin with real problems of social adjustment made significantly higher scores on the intelligence tests than the better adjusted twin. Possibly the very difficulties of adjustment caused Twin A to turn to intellectual preoccupations as a form of compensation (Cases 3A and 3B in Figures 14, 15, and 16).

A boy who never seemed highly motivated, intellectually, earned fair intelligence scores, but had difficulty graduating from high school. He suddenly evidenced strong motivation on his 25-year test, with a large increase in score. A short time before this last test he had struck out on his own for the first time, on a job that was not connected with the family business, and he was belatedly making good (Case 7M in Figure 26).

A girl made phenomenal intellectual growth between 2 and 10 years. During much of this time she was invalided with asthma. She was strongly competitive with an older sibling, and used her enforced rest periods in reading and imaginative games.

A boy who was very slow in both motor and mental development during infancy made rapid intellectual progress after two years: He was highly verbal, but never gained much skill in his motor coordinations. Here the emerging independence of intellectual capacities seems to have permitted their more rapid growth (Case 5M, Figure 18).

One boy was undernourished in infancy, with accompanying retardation in physical maturing and growth, but this seemed unrelated to his relatively steady mental development.

Several children suddenly slowed down in their skeletal maturing, at around 13 years, with subsequent prolonged slow pre-adolescent growth. Two of these, with their later spurts of growth, eventually attained taller statures than had been predicted before 13. In these cases there appeared to be more or less normal delays in starting the steroid phase of growth.

There was at times evidence that a child was "bored" with repetition of the same old mental tests, and the resulting poor effort lowered the score. An adolescent girl, suddenly aware of her plumpness and more curvaceous figure, would diet and lose weight. And of course, severe illnesses often caused temporary slowing down of the physical growth processes.

There are as many more such examples as there are children in the Study. However, there is not space here for them all or even to go into the details to support these brief descriptive statements. But the data are there, and it can be done. The more one goes over these cases, the more evident it becomes that the pattern of growth in each child is unique. We can compare him with his peers, and with his own past history. Often, after the fact, we can find an explanation of the causes of his deviations, some of which seem to be inherent and some environmental. With the accumulation of this kind of information, we can hope gradually to develop general rules that will permit us to make some evaluations and some tentative predictions, and even, in some instances, to correct undesirable deviations.

I hope these evidences of the complexity in the processes of child development will not be too discouraging. On the contrary, they should spur us on to learn more about their intricate interrelationships. Only in this way may we eventually set up optimal conditions for growth and development.

REFERENCES

1. BAYLEY, N. Mental growth during the first three years. *Genet. Psychol. Monogr.*, 1933, 14.
2. BAYLEY, N. Size and body build of adolescents in relation to rate of skeletal maturing. *Child Develpm.*, 1943, 14, 47-90.
3. BAYLEY, N. Skeletal maturing in adolescence as a basis for determining percentage of completed growth. *Child Develpm.*, 1943, 14, 1-46.
4. BAYLEY, N. Consistency and variability in the growth of intelligence from birth to eighteen years. *J. genet. Psychol.*, 1949, 75, 165-196.
5. BAYLEY, N. Some increasing parent-child similarities during the growth of children. *J. educ. Psychol.*, 1954, 45, 1-21.

6. BAYLEY, N. On the growth of intelligence. *Amer. Psychologist,* 1955, 10, 805-818.

7. BAYLEY, N. The XM standards of skeletal maturing. Series of X-rays of hand and knee, for boys and for girls, ages 8 years to maturity. Based on data from the Univer. of California Institute of Child Welfare. Unpublished.

8. BAYLEY, N., & DAVIS, F. C. Growth changes in bodily size and proportions during the first three years: A developmental study of 61 children by repeated measurements. *Biometrika,* 1935, 27, 26-87.

9. BAYLEY, N., & JONES, H. E. Environmental correlates of mental and motor development: A cumulative study from infancy to six years. *Child Develpm.,* 1937, 8, 329-341.

10. BAYLEY, N., & PINNEAU, S. A. Tables for predicting adult height from skeletal age: Revised for use with the Gruelich-Pyle hand standards. *J. Pediat.,* 1952, 40, 423-441.

11. GOLDBERG, MINNIE B. What makes us grow as we do. *J. Amer. med. Women's Ass.,* 1955, 10, 110-116.

12. JONES, H. E., & BAYLEY, N. The Berkeley Growth Study. *Child Develpm.,* 1941, 12, 167-173.

13. MACFARLANE, J. W. Studies in child guidance. I. Methodology of data collection and organization. *Monogr. Soc. Res. Child Develpm.,* 1938, 3, No. 6.

14. OLSON, W. C., & HUGHES, B. O. The concept of organismic age. *J. educ. Res.,* 1942, 35, 525-527.

EXPERIMENTAL RESEARCH IN CHILD PSYCHOLOGY

Boyd R. McCandless and Charles C. Spiker
Iowa Child Welfare Research Station

From the literature of child psychology, and from discussions with child psychologists throughout the country, the writers have obtained the definite impression that child psychologists are concerned and somewhat troubled about the present and future status of their field. In particular, two deficiencies in the area appear to be arousing consternation. The first concern has to do with an alleged deficiency in the quality and quantity of research conducted by child psychologists, and the second with an alleged lack of interest in theory on the part of child psychologists. The writers agree that these two deficiencies exist. Rarely, however, have they observed the two discussed in relation to each other. It seems to them that there is actually a close connection between the type of research that has predominated in the field and the failure of theory to develop. It is the purpose of this paper to attempt to show what this connection is.

First, perhaps, a few comments should be made concerning the essential characteristics of scientific theories. A scientific theory consists of (*a*) a set of clearly defined terms that refer to empirical phenomena; (*b*) a set of basic principles (axioms) that relate many of the concepts to each other; and (*c*) a further set of empirical laws that may be deduced, according to the rules of mathematics and/or formal logic, from the basic principles. For the scientist, explanation consists in the deduction of the phenomenon to be explained from the set of basic principles. In this strict interpretation of scientific theory, it is clear that neither child psychology nor general psychology has thus far met with more than moderate success in theory construction.

There are no logical rules to specify what characteristics the axioms of a theory must have, other than that additional laws must be deducible from them. There are, however, some empirical considerations that are suggested

CHILD DEVELOPMENT, Vol. 27, No. 1 (March, 1956)

by examining the histories of other, more mature sciences. For one thing, the axioms of mature scientific theories have contained concepts that refer to basic variables. The discovery of which variables are basic demands that the scientist isolate, in some manner, the independent variable in such a way that he can discover if and how it affects the dependent variable. The usual method of isolating the independent variables is by means of the scientific experiment. The scientist sets up an experiment in which the independent variable is under his control and manipulation. He then systematically varies the independent variable and determines whether or not, and how, it affects the dependent system. It is, of course, not always possible or necessary to manipulate the independent variable experimentally. The most extensive use of a non-manipulative method of isolating the independent variables can be found in classical astronomy, where none of the variables is subject to the scientist's manipulations. It is interesting to note, however, that Newton's theoretical integration of terrestrial and celestial mechanics made extensive use of Galileo's experimental data on falling bodies, and it is improbable that he would have invented his concept of mass if he had not had the Galileo data available.

While a relevant variable can be isolated without the experimenter's having control over it, the difficulty of this procedure is vastly increased as the number of relevant variables affecting the dependent system is increased. That a very large number of variables affects the behavior of the living organism is a fact that has long been recognized. And if several variables, each of which is a potentially relevant variable, are simultaneously present, the experimenter cannot possibly say which has affected his dependent variable, nor to what degree.

Examination of a list of "independent" variables in child psychology leads one to the conclusions that (*a*) it is often difficult to say which are the dependent and which are the independent variables (e.g., correlations among various tests or other response measures); (*b*) where it is possible to make a distinction, it is often clear that the "independent" variables cannot possibly be either manipulated or isolated from other possibly relevant variables (e.g., chronological age, level of parent education, racial or nationality background, intelligence, socio-economic status, etc.); and (*c*) in many cases, the so-called independent variables were never meant to be anything more than a name for a whole host of potentially relevant variables (e.g., chronological age, sex, school grade, delinquency, etc.). It was undoubtedly in reaction to the selection of such independent variables that Lewin made his persistent pleas for the study of genotypic rather than phenotypic variables in child behavior. For Lewin, genotypic variables were worthy candidates for inclusion in the axioms of theory; phenotypic variables were those that had to be accounted for in terms of the theory.

Another useful consideration in the construction of a theory arises from the recognition that somewhere among the axioms of the theory there must be some that instruct the scientist about the analysis of more complex phenomena than he has yet encountered experimentally. The scope of a theory—that is, the number and variety of phenomena that it can explain—is closely dependent on the existence of these composition laws among its axioms. Some of the best, and simplest, examples of such laws in psychology can be found in current conditioning theories of learning, where there are laws that specify how conditioning principles can be made to apply in situations in which conditioning and extinction occur at the same time, in which two or more stimuli conditioned to the same response are presented simultaneously, in which two or more stimuli evoking incompatible responses are given together, and so on. Such laws are almost always discovered through series of experiments designed to discover the limits of previous generalizations and to isolate the conditions under which previous generalizations actually hold.

It is indeed rare that one finds empirical laws in child psychology that show any promise of successful use as composition laws in the construction of theory. It seems doubtful, for example, that anyone expects the concept of chronological age to be among the basic concepts in any theory of behavior. Rather, we recognize that changes in behavior associated with chronological age are to be explained in terms of amount of previous learning experiences; of physiological and anatomical changes; of changes in the nature of social rewards, punishments, and status; and of other variables that are also more or less incidentally related to time since birth. In spite of this, the research literature reports more relationships of behavior to age than to any other variable except, possibly, intelligence.

It seems to the writers that the fundamental need of child psychology—the development of promising theories—has been greatly handicapped by child psychologists' tendencies to avoid experimental research. In stating this, there is no implication that laws discovered in nonexperimental research are useless; no responsible scientist scoffs at well-established laws. Rather, the objection is to the imbalance. Child psychologists may well look to general psychology for an object lesson. As limited as the theories in general psychology are, compared to what they will undoubtedly be some day, it is not difficult to see the tremendous strides that the field has made in theory construction during the last 25 years. Examination of the empirical bases of these theories shows that they have been obtained from experimentation in which relevant variables have been isolated and controlled.

The writers were trained and have worked as child psychologists, and believe they understand the factors that have led child psychologists away from experimental research. Carefully planned research requires time—time which many child psychologists do not have because the tendency of their employers is to require heavy teaching and service loads of them. Experimental research worthy of publication often requires numerous unpublishable preliminary experiments, which require still more time. On the other hand, it has been the experience of the writers that it is a relatively simple matter for an ordinary insightful research worker to take several tests, administer them to a group of subjects, obtain the intercorrelations, and produce some correlation coefficients that are statistically significant. In this same connection, it is usually easier to arrange with school administrators to obtain subjects for two or three hours of time from total classes than it is to arrange to have the subjects on an individual basis for a half-an-hour or so, even more so when a frequently illusory appearance of immediate application of results exists.

Another important factor is precedent. We have in mind not only a social or professional reluctance to stray from a trodden path, but also the additional work that is required to do so. Only a few fruitful experimental apparatus and procedures have been worked out for use with children and many of these have been used so rarely that little useful normative information is available. Thus, the child psychologist who is inclined to be experimental must first invent appropriate apparatus and develop workable procedures and instructions. A similar difficulty arises when the potential experimenter attempts to select a problem to investigate. Few of the textbooks and articles in child psychology can be said to be rich sources for experimental hypotheses. On the contrary, most of them are preoccupied with the application of available knowledge. The problems that suggest themselves to the research worker tend to be concerned with immediate, practical situations. This tendency is reinforced by the nature of the duties of the typical child psychologist. He has usually been hired primarily to teach and to perform service. The types of problems that lend themselves to experimental investigation, under such employment and service exigencies, do not seem "big" enough to him, i.e., he can see no immediate prospect of their helping practically to serve society.

All the reasons for our failure to have produced a wealth of experimental research, unfortunately, are not as socially acceptable as those just mentioned. The fact is that child psychologists as a group, not unlike clinical psychologists as a group, have not been trained to do experimental research. The literature testifies to our naivete with respect to statistical techniques, the philosophy of science, research methods, and understanding of theory construction. Our statistical analyses are often outmoded and otherwise inadequate; our reasoning is sometimes confused and our concepts are often vaguely defined; our control groups are all too frequently poorly conceived; our notions of what constitutes scientific theory are most unsophisticated.

In order to improve our scientific status, there are several concrete steps that we may take. First, we should stop marking time, waiting for a brilliant theorist to appear who will organize our data into a satisfactory theory. Only a mind sufficiently clairvoyant to intuit the missing basic laws could possibly integrate the data we have available. Second, we can well look to the relatively more sophisticated field of general psychology for suggestions concerning apparatus, procedures, and hypotheses. There are many techniques that have been used successfully with adults and infrahuman organisms that can be relatively easily modified for use with younger children. It is the writers' belief, too, that current theories in general psychology are replete with hypotheses for the test of which children are not only satisfactory subjects, but *ideal* subjects. Those of us who have dealt with both children and adults know, for example, that the purpose of an experiment is more easily disguised from children than from college sophomores; that learning problems for children can be much simpler (and perhaps the laws discovered more basic) than corresponding problems for adults; that the learning and functions of language can be more fruitfully studied with children than with adults; that children are the ideal subjects for the study of imitation learning; that children can be more easily excited, frustrated, and angered in an experimental situation than can adults; and so on. Several recent experiments with children as subjects have been among the most widely quoted in psychological literature. It is not to the credit of our profession that many of these have been conducted by persons not formally trained in child psychology.

Third, while some of us can, and should, deliberately shirk our applied, clinical responsibilities, there is probably little that most of us can do about the applied and teaching demands that our positions place upon our time. But those of us who are in a position to do so can spend some of our energies in placing well-trained students in positions that both require and allow time for research. Evidence indicates that there is a supply of relatively well-paid positions of this type at the present time. Furthermore, there is likely to be an increase in the number of such positions, since many institutions have become aware that they are not producing psychological research with children and that a fruitful research field may be opened up if they can find properly trained personnel.

Fourth, the availability of good research positions will not help our field if we do not have properly trained personnel to fill them. While there are limits to the extent to which those of us who are several years out of graduate school can improve our personal research competence, there is much that we can do for our students. We can provide them with graduate curricula that include adequate training in statistics, considerable work in the philosophy of science, and thorough familiarity with current theories and with both current and past research in general psychology.

The writers also think that there are more economical ways of achieving familiarity with the conventional content of child psychology and development than have been employed by training centers traditionally. Lectures can be pointed toward organizing and evaluating, rather than endlessly summarizing factual material. Theoretical hunches and fractional ordering concepts about the material can serve as both an interesting and an integrating framework for students, who can then read what portion of the vast accumulation of normative and developmental material they or their instructors believe to be essential. From such an instructional approach, research ideas may even germinate. The writers suspect that less emphasis on applied duties may profitably (or at least without harm) be given to graduates in training, and that instead of being drilled and practicumed for every imaginable future professional contingency, they may be trained to *think* and to react flexibly and adaptively instead of as well drilled automatons. We feel it is vital to the field that we instill in our students the *fundamental* importance of doing experimental research with children.

Given properly trained personnel who conduct sound experimental research on problems oriented toward theory construction, our contemporary concern for the deficiency in theoretical integration within our field can be constructively put to use and eventually allayed.

CHILD DEVELOPMENT

Vol. 27
June, 1956
No. 2

Editor
WILLIAM E. MARTIN
Purdue University

Consulting Editors

REGINALD MACGREGOR ARCHIBALD
Rockefeller Institute for Medical Research

JOSEPH H. BRITTON
Pennsylvania State University

SIDNEY W. BIJOU
University of Washington

VAUGHN J. CRANDALL
Fels Research Institute

ERCEL S. EPPRIGHT
Iowa State College

C. GLEN KING
Nutrition Foundation

BENJAMIN PASAMANICK
Ohio State University

ANSELM STRAUSS
University of Chicago

RALPH L. WITHERSPOON
Florida State University

Managing Edit
KATE HOFFM

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Edn.
Dated
Acc. No

Printed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

EDITORIAL NOTE

This issue of *Child Development* is double-sized, since it includes a Supplement on the Chicago Conference on Child Development held on October 26-29, 1955, under the joint sponsorship of the Elizabeth McCormick Memorial Fund and the Society for Research in Child Development with Donald Brieland, Director of the Fund, and William W. Greulich, President of the Society, as co-chairmen.

The stated purposes of the Conference were: (*a*) to give an opportunity for participating institutions to summarize their program interests and research-in-progress; (*b*) to indicate data which they have available which may be of interest to other people in the field of child development; (*c*) to explore the feasibility of pooling certain data for research purposes; (*d*) to exchange research ideas of the people represented, including consideration of problems in which they are interested but with which they do not plan to deal themselves.

Prior to the Conference, representatives of each participating university or research group prepared and circulated to all participants written summary descriptions of research-in-progress, data available and publications. Summaries were prepared by the following:

Nancy Bayley, National Institute of Mental Health, National Institutes of Health, Public Health Service, U.S. Department of Health, Education and Welfare
Dale B. Harris, Institute of Child Welfare, University of Minnesota
Harold E. Jones, Institute of Child Welfare, University of California
Boyd R. McCandless, Iowa Child Welfare Research Station, State University of Iowa
Jean Walker Macfarlane, Guidance Study, Institute of Child Welfare, University of California
Lester W. Sontag, Fels Research Institute, Antioch College
John W. M. Whiting, Laboratory of Human Development, Harvard University

These summaries provided the basis for three discussion sessions on Hypotheses, Methodology and Implications for Theory in Current Research which were led respectively by John D. Benjamin, Denver Child Research Council; Roger G. Barker, Department of Psychology, University of Kansas; and Alfred L. Baldwin, Department of Child Development and Family

Relationships, Cornell University. These papers and the ensuing discussions do not appear in the Supplement.

A second set of papers dealing with the research concerns and activities of individuals was also prepared and distributed prior to the conference. These papers, together with two delivered at the Conference by John E. Anderson and Harold E. Jones, constitute the Supplement. They have been revised and abridged to meet space requirements. A longer paper prepared for the Conference by Helen L. Koch, Department of Psychology, University of Chicago, will appear in a later issue of *Child Development.*

The Elizabeth McCormick Memorial Fund underwrote the costs of the Conference and made possible the publication of the Supplement. To the Fund and its Director go the gratitude of the members of the Society for Research in Child Development and the readers of *Child Development.*

WILLIAM E. MARTIN
Editor

THE STIMULUS GENERALIZATION GRADIENT AS A FUNCTION OF THE INTENSITY OF STIMULUS LIGHTS

CHARLES C. SPIKER[1]
Iowa Child Welfare Research Station

The experiment reported here is the second in a series designed to develop procedures and apparatus appropriate for the study of the generalization behavior of children. In the first experiment (10), the amount of generalization was shown to be an increasing function of the number of reinforcements given on the training stimulus. The present experiment is concerned with the hypothesis that the shape of the generalization gradient, for stimulus intensity dimensions, is dependent upon the absolute values of the stimuli selected as well as upon the degree to which the stimuli are separated on the stimulus dimension.

Several investigators have found that if the subject (*S*) is conditioned to an intense stimulus and is then tested on stimuli of lesser intensity, the generalization gradient is steeper than if the *S* is conditioned to a weak stimulus and tested on more intense stimuli. Hovland (3) reported this finding with the GSR for human *S*s using auditory stimuli; Brown (1) and Perkins (8) obtained similar findings for an instrumental approach response of rats to lights differing in brightness; and Grice and Saltz (2) demonstrated the phenomenon in rats' approach responses to stimuli differing in size. Hull (4) attempted to explain some of these results in terms of his principle of stimulus dynamism. Although the details of his explanation are not given here, it involves the assumption that the intensity of a stimulus contributes to the general motivational level of *S*, and that the motivational factors multiply with the generalized habits. Thus, given two stimuli (S_1 and S_2) that are equally separated from the conditioned stimulus, S_1 being more intense and S_2 less intense than the conditioned

---

[1] The writer is indebted to Miss Kathryn Norcross and Mr. Lewis P. Lipsitt for aid in collecting the data and for valuable suggestions regarding procedural matters.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

stimulus, S_1 will have a greater reaction potential than will S_2, The deductive consequence of these assumptions, and others in the system, is that the generalization gradient will be steeper when the test is made with less intense stimuli than when made with more intense stimuli.

In this experiment, the tendency to respond to the conditioned stimulus was established by giving differential reinforcement during training, in which *S* was reinforced for responding to the conditioned stimulus and was nonreinforced for responding to another stimulus that was different from the conditioned stimulus in terms of one characteristic (hue), but the same with respect to the dimension being investigated (brightness). Subsequently, *S*'s tendency to respond to other stimuli, differing from the training stimulus only in brightness, was determined in a generalization test. Two levels of reinforcement during training were also used in an attempt to confirm the findings of the previous experiment (10).

METHOD

Subjects and Experimental Design

The *S*s were 60 children from the Iowa Child Welfare Research Station preschool laboratories. They ranged in age from 3 years, 9 months, to 5 years, 8 months. Most *S*s had been used in a previous study of generalization (10) in which the present apparatus was used, but in which the stimuli differed in hue rather than in brightness.

The experimental design was such that one-half the *S*s were randomly assigned to a group that was trained to respond to a bright light and then tested on dimmer lights. The other half were assigned to a group trained to respond to a dim light and tested on brighter lights. One-half of each of these groups was randomly assigned to a group that received 12 presentations of the postive (white) light and 12 presentations of the negative (blue) light during reinforcement training. The other halves of the two groups received 24 presentations of each of the positive and negative stimuli. For convenience, these four subgroups will be referred to as B-12 (bright light, 12 presentations of each stimulus); B-24 (bright, 24 presentations); D-12 (dim, 12 presentations); and D-24 (dim, 24 presentations).

Apparatus

The experiment was conducted in a semi-darkened experimental room with an adjoining observation room. The apparatus, which has been described elsewhere in detail (10), was placed in the window between the two rooms. *S* operated the apparatus from the experimental room; the experimenter (*E*) operated the controls and recorded *S*'s responses from the observation booth. S faced a gray metal panel, 20 in. wide and 12 in. high; 3 in. from the top was centered an aperture, 2 in. in diameter, covered with flashed opal glass. Below and to the left of the aperture was a transparent container into which marbles were ejected. A metal rod with a hand grip extended from the panel below and to the right of the aperture. Directly below the apparatus was a wooden stand, 2 ft. square and 8 in. high, on which *S* stood during the experiment.

The apparatus was enclosed in a metal cabinet, 12 in. high, 20 in. wide, and 12 in. deep. Approximately 7 in. behind the aperture were six 75-watt projection lamps, arranged so that their beams focused on the aperture. Each lamp was enclosed in a reflecting container on the front of which 2 in. square wooden diaphragms or glass color filters could be mounted. A stepping relay controlled the current to the lamps and the relay was in turn controlled by two interval timers. One of the timers controlled the presentation interval and the other controlled the interval between presentations. A multi-contact selector switch was used to determine which stimuli would be presented and the order of their presentation.

If a stimulus lamp was on at the time the response occurred, the response was recorded on one of six electrical impulse counters. Each counter was associated with one stimulus, so that the responses to each stimulus were separately recorded. Approximately 8 pounds of pressure were required to move the lever the ¾ in. necessary to activate the counters and the marble ejector. The marble ejector was activated by a response only on those trials during which the reinforced stimulus was presented.

Stimuli

The stimuli were lights, different brightnesses being obtained by placing wooden diaphragms between the lamps and the stimulus aperture. The brightnesses of four stimuli, measured at the aperture with a McBeth Illuminometer, were approximately 2.5, 25, 80, and 250 ft.-candles. The difference between the common logarithms of 2.5 and 25, and of 25 and 250 is 1.0. The 80 ft.-candle stimulus is one-half a logarithm unit between the 25 and 250 ft.-candle stimuli. One-half the *Ss* were trained with the 250 stimulus, while the other half were trained with the 2.5 stimulus. Two additional stimuli were constructed by placing a blue glass color filter over the diaphragms for the 2.5 and 250 ft.-candle stimuli. This produced two pairs of stimuli, the members of each pair approximately equal in brightness but differing in hue. The blue stimuli were used as the nonreinforced (negative) stimuli in the training sessions.

Preliminary Training.

S was brought to the experimental room, shown a variety of dime-store toys that were spread out on a small table, and was allowed to choose the toy that he wished to have. *E* informed him that he must buy the toy with marbles and that *E* would show him how to get the necessary marbles. The apparatus was turned on and set to present the training stimulus series. This series consisted of an equal number of separate presentations of the white (positive) and of the blue (negative) stimuli. Groups B-12 and B-24 were presented with a white and blue light of approximately 250 ft.-candles; Groups D-12 and D-24 received a white and a blue light of approximately 2.5 ft.-candles. Throughout the experiment, the presentation period was 3 seconds and the inter-presentation interval was 2½ seconds.

E first demonstrated the apparatus by pulling the lever several times while the training light was on and pointing to the container into which the marbles were being ejected. When the blue light came on, *E* pulled the lever a few times to demonstrate that no marbles were being ejected. *S* was then allowed to pull the lever to obtain marbles. The preliminary training was continued until *S* learned to pull more than once when the training light was on and failed to pull the lever between stimulus presentations. As soon as *S* had met these criteria, *E* stopped the apparatus and returned the stimulus presentation device to the starting position in preparation for the reinforcement training.

Reinforcement Training

E returned *S* to his position before the apparatus, and informed him that *E* was going into the next room to start the apparatus again and that *S* would be informed when he had obtained enough marbles to buy the toy. *E* went into the observation booth and started the apparatus. The reinforcement stimulus series was the same one given during the preliminary training. In each successive block of six presentations, there were three each of the positive and negative stimuli. Neither stimulus was presented more than twice in succession. Groups B-12 and D-12 were given 12 presentations of each the white and the blue lights and Groups B-24 and D-24 were given 24 presentations of each. *E* recorded on prepared sheets the number of responses *S* made to each stimulus presentation.

Generalization Test

Following the last trial on the reinforcement training, *E* turned the selector switch so that the generalization stimulus series was presented. The time between the last reinforcement training trial and the first generalization test trial was eight seconds. The generalization series was introduced with no other changes in the procedure.

The series consisted of two cycles, each cycle having seven presentations of the positive stimulus with each response being reinforced, and five nonreinforced presentations of the test stimuli. The test stimuli were presented on trials 2, 4, 7, 9, and 12 of the cycle. The test stimuli included the positive stimulus, the negative stimulus, and each of the other three stimuli differing from the positive stimulus in brightness. The test stimuli were presented in counterbalanced order so that within each group each test stimulus appeared in each position an equal number of times. For each *S*, the second cycle was the same as his first.

RESULTS

Responses to the Reinforced Stimulus during the Generalization Series

An analysis was first made of the number of responses to the reinforced training stimulus during the generalization test. This analysis should reveal the differential effects on learning, if any, of the different conditions that prevailed during the acquisition stages. A Lindquist Type III analysis (6, p. 281) was conducted, in which the main effects were the number of reinforced trials during training, the brightness of the conditioned stimulus, and successive three-trial blocks for the first 12 of the 14 reinforced trials of the generalization test. The results of the analysis need not be presented in detail, since only the trial-block effect proved significant at less than the 5 per cent level. This significant effect reflects the fact that the number of responses to the reinforced conditioned stimulus tended to decrease with an increase in the number of previous test trials. This decrease is probably due to the generalized effects of nonreinforcement of the test stimuli.

This analysis makes tenable the hypothesis that the four groups did not differ significantly in performance to the conditioned stimulus as a function of the different conditions that obtained during training. This appears to be true both for the brightness of the conditioned stimulus and for the number of reinforcements received during training.

Responses to the Test Stimuli for the Complete Generalization Test

Table 1 gives the mean number of responses made to each test stimulus for the first and second cycles of the generalization test series. An examina-

TABLE 1

MEAN NUMBER OF RESPONSES TO EACH TEST STIMULUS

Test Stimulus	*First Cycle* Training Reinforcements 12	24	*Mean*	*Second Cycle* Training Reinforcements 12	24	*Mean*
	BRIGHT-TRAINED GROUPS					
250	2.40	3.33	2.87	2.00	2.40	2.20
80	1.13	1.60	1.37	1.13	0.93	1.03
25	0.80	1.13	0.97	1.00	0.67	0.83
2.5	0.80	0.53	0.67	0.33	0.53	0.43
Mean	1.28	1.65	1.47	1.12	1.13	1.12
	DIM-TRAINED GROUPS					
2.5	2.33	1.80	2.07	1.93	2.27	2.10
25	1.40	1.87	1.63	1.60	1.07	1.33
80	1.40	1.27	1.33	0.87	1.13	1.00
250	1.07	1.47	1.27	0.67	1.00	0.83
Mean	1.55	1.60	1.58	1.27	1.37	1.32

tion of this table reveals (*a*) a somewhat greater mean number of responses for the Groups B-24 and D-24 than for D-12 and B-12; (*b*) a greater mean number of responses for all groups combined during the first cycle than during the second cycle; (*c*) a slightly greater mean number of responses to the test stimuli for Groups D-24 and D-12 than for Groups B-24 and B-12; (*d*) a decrease, for all groups combined, in the number of responses to the test stimuli as the difference between training and test stimuli increases; and (*e*) a relatively greater decrease in these responses for Groups B-12 and B-24 than for Groups D-12 and D-24. The appropriate analysis for testing the statistical significance of these observations is an extension of Lindquist's Type VI design (6, p. 292). The main factors in the analysis are the number of reinforcements during training, the brightness of the training stimulus, the different test stimuli, and the cycles of the generalization test.

In order to analyze the data, the stimulus dimension was reversed for *Ss* trained to the dim stimulus. That is, the responses made by these *Ss* to the 2.5 ft.-candle stimulus, their training stimulus, were compared with the

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF RESPONSES TO EACH TEST STIMULUS PRESENTATION

Source	*df*	*Mean Square*	*F*	*p*
Subjects	59			
Reinforcement (R)	1	1.88		
Training Stimulus (T)	1	4.01		
R × T	1	0.40		
error (b)	56	4.14		
Within Subjects	300			
Test Stimuli (S)	2	73.51	41.53	<.001
Cycles (C)	1	7.51	11.73	<.005
S × C	2	0.12		
S × T	2	9.88	5.58	<.005
S × R	2	0.68		
C × T	1	0.28		
C × R	1	0.54		
error (w)	280			
error$_1$ (w)	56	0.64		
error$_2$ (w)	112	1.77		
error$_3$ (w)	112	1.06		
Total	359			

NOTE.—Higher order interactions have been omitted, since none was significant at less than the 10 per cent level. The stimulus dimension was reversed for *Ss* trained to the dim stimulus, and responses of all *Ss* to the 80 ft.-candle stimulus have been omitted.

responses made by the bright-trained *Ss* to their 250 ft.-candle training stimulus. Responses made by the dim-trained group to the 25 ft.-candle stimulu were compared with those made by the bright-trained group to the sam stimulus; and responses made by the dim-trained group to the 250 ft. candle stimulus were compared with those made by the bright-traine group to the 2.5 ft.-candle stimulus. By omitting the responses of bot groups to the 80 ft.-candle stimulus, there are equal logarithmic steps between the training and test stimuli for both groups.

A summary of the analysis of variance is presented in Table 2. Th triple and quadruple interactions have been omitted from the table sinc

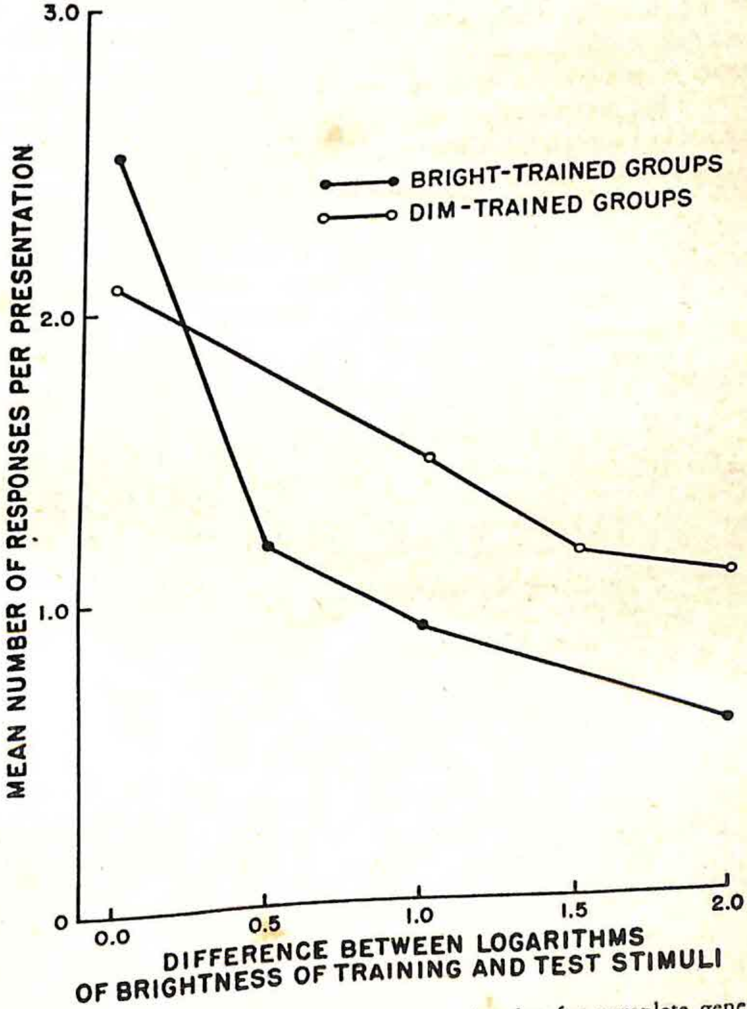


FIGURE 1—Mean number of responses per test stimulus for complete generalization series.

none was significant at less than the 10 per cent level. As the table indicates, neither the reinforcement factor (R) nor the brightness of the training stimulus (T) proved to be significant. The significant cycle effect indicates the reduction in the number of responses made to the test stimuli from the first to the second cycle. The significant test stimulus effect indicates that the generalization gradient for the combined groups cannot be attributed to chance. The significant S × T interaction reflects the fact that the generalization gradient for *S*s trained to the bright stimulus is significantly steeper than the gradient for *S*s trained to the dim stimulus. This latter finding is illustrated in Figure 1, where the means of the numbers of responses per test stimulus presentation, for the pooled bright-trained and pooled dim-trained groups, are plotted against the difference between the common logarithms of the training and test stimuli. The mean number of responses to the 80 ft.-candle stimulus is also included for each group. It is apparent from Figure 1 that the gradient for the bright-trained *S*s is steeper than that for the dim-trained.

In order to establish the existence of a generalization gradient for the bright-trained and the dim-trained groups, Lindquist's Type I analysis (6, p. 267) was made separately for these two groups. In this analysis, responses to the 80 ft.-candle stimulus were included. The response measure was the mean number of responses to each test stimulus for the entire generalization test series (disregarding cycles) and the level of reinforcement during training was ignored. Table 3 presents a summary of these analyses. The effect of primary interest here is that for test stimuli, since a significant test stimulus effect indicates a significant generalization gradient. For both groups, the generalization gradient proves to be highly significant ($p < .001$). Table 4 gives the probabilities resulting from *t* tests of the significance

TABLE 3

SUMMARY OF ANALYSIS OF VARIANCE OF THE NUMBER OF RESPONSES TO THE TEST STIMULI FOR THE BRIGHT-TRAINED AND THE DIM-TRAINED GROUPS SEPARATELY

Source	*df*	*Mean Square*	*F*	*p*
	BRIGHT-TRAINED GROUPS			
Test Stimuli (S)	3	98.48	40.53	<.001
Subjects (I)	29	9.46		
S × I	87	2.43		
	DIM-TRAINED GROUPS			
Test Stimuli (S)	3	25.69	7.56	<.001
Subjects (I)	29	12.02		
S × I	87	3.40		

NOTE.—This analysis includes the responses to the 80 ft.-candle stimulus.

of the differences between the means for pairs of stimuli for the two groups separately. It appears that the significance of the gradient is primarily due to differences between responses to the training stimulus (under test conditions) and the other test stimuli. For the bright-trained group, however, there is a significant difference ($p < .005$) between the means for the 80 ft.-candle and the 2.5 ft.-candle stimuli.

TABLE 4

PROBABILITIES OBTAINED FROM *t* TESTS COMPARING THE MEANS FOR PAIRS OF TEST STIMULI FOR THE BRIGHT-TRAINED AND DIM-TRAINED GROUPS SEPARATELY

Bright-Trained Group		Dim-Trained Group	
Stimulus Pairs	*p*	*Stimulus Pairs*	*p*
250 vs. 80	<.001	2.5 vs. 25	<.02
250 vs. 25	<.001	2.5 vs. 80	<.001
250 vs. 2.5	<.001	2.5 vs. 250	<.001
80 vs. 25	<.20	25 vs. 80	<.20
80 vs. 2.5	<.005	25 vs. 250	<.10
25 vs. 2.5	<.10	80 vs. 250	<.70

Responses to the Initial Test Stimulus Presentation

The failure to find a significant effect for reinforcement during training, contrary to the results of other studies, suggested that the additional reinforcements given during the generalization test may have obliterated any differences that existed initially between the two reinforcement groups. Accordingly, the number of responses made to the first test stimulus presented was obtained for each *S*. There were three *Ss* in each of the four main groups that received the same test stimulus on the first presentation. Since one of these subgroups received the negative stimulus first and another received the 80 ft.-candle stimulus, neither of which are of concern here, there remained 36 *Ss*, nine in each group, who could be used in this analysis. The means of these cases, each cell based on three *Ss*, are presented in Table 5. Trends corresponding to those in Table 1 may be observed here. A Lindquist Type III analysis (6, p. 281) was made of these data and the summary is presented in Table 6. The test stimulus variable is again highly significant. The data show a significant ($p < .025$) reinforcement effect, and a significant ($p < .05$) interaction between reinforcements and test stimuli. The interaction between brightness of the training stimulus and the test stimuli is not as pronounced as in the first analysis, although the pattern of the corresponding means is the same. In Figure 2, the means for *Ss* receiving 12 reinforcements and the means for those receiving 24 reinforcements are plotted as a function of the test stimuli. It is apparent that the reinforcement effect is due to the greater number of generalized responses by the 24 reinforcement *S*s and that the interaction is due to the steeper slope of the gradient for these *S*s. A *t* test of the simple effects shows that the group receiving 24 reinforcements made significantly more responses to the first and second test stimuli ($p < .05$). This analysis thus indicates that the greater number of reinforcements during the training resulted in significantly greater generalization to the stimuli near to the training stimulus, at least during the early stages of the generalization test.

TABLE 5

MEAN NUMBER OF RESPONSES TO FIRST TEST STIMULUS PRESENTED THE SUBJECTS

		Reinforcements During Training		
Training Stimulus	*Test Stimulus*	*12*	*24*	*Mean*
250	250	1.33	3.00	2.17
	25	0.33	1.33	0.83
	2.5	0.67	0.33	0.53
	Mean	0.78	1.56	1.17
2.5	2.5	1.67	2.33	2.00
	25	1.33	2.33	1.83
	250	0.67	0.33	0.50
	Mean	1.22	1.67	1.44

TABLE 6

ANALYSIS OF VARIANCE OF THE MEAN NUMBER OF RESPONSES TO FIRST TEST STIMULUS PRESENTED TO SUBJECTS

Source	*df*	*Mean Square*	*F*	*p*
Training Stimulus (T)	1	0.70		
Reinforcements (R)	1	3.36	5.89	$<.025$
Test Stimuli (S)	2	7.53	13.21	$<.001$
T × R	1	0.25		
T × S	2	1.19	2.09	$<.20$
R × S	2	2.02	3.54	$<.05$
T × R × S	2	0.25		
Within Cells	24	0.57		
Total	35			

NOTE.—The stimulus dimension has been reversed for *S*s trained to the dim stimulus, and *S*s given the 80 ft.-candle stimulus first have been omitted.

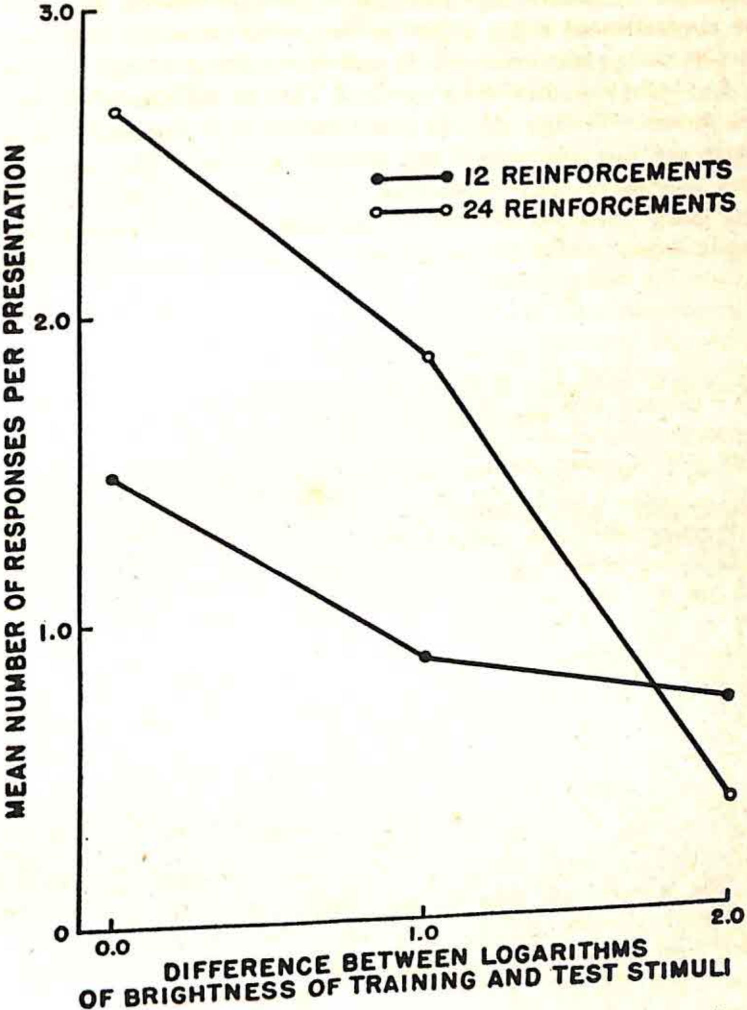


Figure 2—Mean number of responses to first test stimulus presentation.

Discussion

Analysis of the number of responses made to the test stimuli during the generalization test reveals a highly significant stimulus generalization gradient for all *Ss* combined. That this "group gradient" is not merely an artifact of combining individual step-wise gradients is demonstrated by the fact that 80 per cent of all *Ss* responded fewer times to the most distant test stimulus than to the nonreinforced training stimulus and 60 per cent responded fewer times to the most distant test stimulus than they did to the test stimulus nearest the training stimulus.

Analysis also indicated that the gradient for *S*s trained to the bright stimulus is significantly steeper than that for *S*s trained to the dim stimulus. This finding is in agreement with that of Grice and Saltz (2) for the size dimension, with that of Brown (1) and Perkins (8) for the brightness dimension, and with Hovland's (3) for loudness. It is also in general agreement with the prediction made by Hull (4) on the basis of the principle of stimulus dynamism. Further analysis revealed that the generalization gradient for the dim-trained *S*s is statistically significant, indicating that a generalization gradient can be demonstrated even when stimulus intensity operates in the opposite direction.

The significant effect of cycles—representing repeated presentations of nonreinforced test stimuli—is also one of the generalization phenomena frequently observed. This principle is one of the main axioms in stimulus-response theories that are designed to account for the various phenomena in discrimination learning (5, 9).

Although the effects of the number of reinforcements during training failed to achieve significance in the above described analysis, a significant reinforcement effect was obtained when only the responses to the first stimulus presentation were analyzed. In addition, a significant interaction between test stimuli and number of reinforcements was found, indicating a steeper gradient for the *S*s receiving 24 reinforcements than for the group receiving only 12. This result may also be derived from Hull's formulations, and has been empirically demonstrated for the instrumental approach response of rats (7).

The results of this experiment indicate that the apparatus and procedures used offer a fruitful and practical means of investigating generalization phenomena with child *S*s. The method has been tested on hypotheses derived from the results obtained with more conventional methods and subjects. The hypotheses that were verified deal with the establishment of a gradient of stimulus generalization; the effects of reinforcement on the amount and shape of this gradient; the effects of repeated nonreinforced stimulus presentations on the magnitude of generalization; and the effect of stimulus intensity on the shape of the gradient. It now appears that the method warrants use in the study of some of the generalization phenomena which may best be studied with children; e.g., the effect on the shape of the gradient of prior name-learning experience with the stimuli; the effects on generalization behavior of previous training to generalize or to discriminate; and the relationship between level of maturity (i.e., chronological or mental age) and the magnitude, range, and steepness of the generalization gradient.

SUMMARY

This experiment is primarily concerned with the study of the effects on generalization of the number of reinforcements with the conditioned stim-

ulus and with the effects of the absolute brightness of the stimuli on the magnitude and shape of the generalization gradient. Sixty preschool *Ss* were given differential reinforcement to a white (positive) and a blue (negative) light. For one-half the *Ss*, the positive stimulus was the brightest of four stimuli differing from each other in brightness; for the other half, it was the dimmest. One-half of each these groups was given 12 presentations and the other half was given 24 presentations of each the positive and negative stimuli. The response was the repeated pulling of a lever for marbles during a 3-second presentation of the conditioned stimulus, and the response measure was the number of such responses that occurred during the 3-second periods. Immediately following training, the *Ss* were tested, without reinforcement, on each of the four stimuli differing in brightness. The main findings were (*a*) a statistically significant generalization gradient for all groups combined; (*b*) a steeper gradient for the bright-trained than for the dim-trained *Ss*, and (*c*) a greater number of generalized responses for *Ss* given 24 reinforced training trials than for *Ss* given only 12, and a steeper gradient for the former than for the latter.

References

1. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, 33, 209-226.
2. Grice, R. G., & Saltz, E. The generalization of an instrumental response to stimuli varying in the size dimension. *J. exp. Psychol.*, 1950, 40, 702-708.
3. Hovland, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensity of tone. *J. genet. Psychol.*, 1937, 51, 279-291.
4. Hull, C. L. Stimulus intensity dynamism (V) and stimulus generalization. *Psychol. Rev.*, 1949, 56, 67-76.
5. Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
6. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton-Mifflin, 1953.
7. Margolius, G. Stimulus generalization of an instrumental response as a function of the number of reinforced trials. *J. exp. Psychol.*, 1955, 49, 105-111.
8. Perkins, C. C., Jr. The relationship between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, 46, 225-231.

9. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

10. Spiker, C. C. The effects of number of reinforcements on the strength of a generalized instrumental response. *Child Develpm.*, 1956, 27, 37-44.

THE EFFECTS OF DIFFERENTIAL REINFORCEMENT AND MOTIVATION ON PREDICTION RESPONSES OF CHILDREN[1]

SIDNEY ROSENBLUM[2]
Iowa Child Welfare Research Station

Since the early 1930's the effects on behavior of varied patterns of partial and continuous reinforcement have received steadily increasing attention from psychologists. Although new and ingenious procedures are being devised for the study of this problem at the human level, the major emphasis continues to be on animal experimentation. A survey of partial reinforcement literature by Jenkins and Stanley (6) reveals that of the 30 studies reported during the years 1939-1949 few used humans as subjects. No significant change in this tendency is apparent from a review of the published research of the past five years. This fact has concerned some psychologists who have argued seriously for the greater utilization of human subjects in the testing of various aspects of learning theory. It is their contention that although constructs and laws coming from the animal laboratory have been shown to predict the behavior of infrahuman organisms with accuracy and usefulness, the time has come for a "validation" of some of these principles at the human level.

Although studies with humans in the area of partial reinforcement have supported, for the most part, results of animal research, a survey of the literature has disclosed two general characteristics of such experiments. First,

---

[1] This article is based on a dissertation submitted to the Iowa Child Welfare Research Station, State University of Iowa, in partial fulfillment of the requirements for the Ph.D. degree, June, 1953. The writer is indebted to Dr. B. R. McCandless under whose direction the research was performed, and to Drs. C. C. Spiker, A. Castaneda, I. E. Farber, and H. Bechtoldt for their helpful criticisms and suggestions.

[2] Now at the Wayne County Training School, Northville, Michigan.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

very few have used the normal child as a subject. Apart from the fact that children can be conceived loosely as representing an "intermediary level" between inarticulate organisms and the ubiquitous college undergraduate, and, as such, constitute another valuable source of data, their use in learning studies seems vital for the development of a systematic approach in the field of child psychology, an area characterized until somewhat recently by a notable lack of sound, theoretically oriented research (7). Second, virtually no studies have been reported in which the effects of motivation have been intentionally introduced and systematically varied. Such situations as have been used [e.g., conditioning of the eyelid response (5), lever-pulling (3), and making predictions about the occurrence or nonoccurrence of a light (4)] would appear to lack "importance" to the subject, in the sense that nothing has been done to make them "ego-involving."

The present study was designed to investigate the effects on children's behavior of various reinforcement patterns in combination with motivational factors. Specifically, it concerns the effects of three schedules of reinforcement (100%, 50%, and 0%) at two levels of motivation (high and low) on the acquisition and extinction of a "certainty of prediction" response (CPR), that is, the degree of certainty with which a child makes a prediction concerning the occurrence of a success experience for himself.

The investigation can be conceived as having five stages. Stage one consists of determining an initial response measure (trial 1). During stage two (trials 1 to 10) changes in the CPR under high and low motivation conditions during 100%, 50%, and 0% reinforcement are determined. Stage three involves observation of the CPR at the end of trial 10. During stage four (trials 11 to 20) the effects of a 0% reinforcement series on the CPR are measured. And stage five (trial 20) consists of obtaining the terminal CPR.

The following hypotheses were formulated for testing:

1. Increase in mean CPR's during trials 1 to 10 would be significantly greater for the 100% group than either the 50% or 0% groups, with the 50% group evincing reliably greater increments than the 0% group.

2. Decrease in mean CPR's during trials 11 to 20 would be significantly greater for the 0% group than either the 100% or 50% groups, with the 100% group demonstrating reliably greater decrements than the 50% group.

3. Increase in mean CPR's during trials 1 to 10 would be greater for high in comparison to low motivation groups, except for the 0% group where no significant differences would manifest themselves.

4. Decrease in mean CPR's from trials 11 to 20 would be greater for low in comparison to high motivation groups, except for the 0% group where no significant differences would manifest themselves.

It should be noted with regard to the motivational variables that there is no systematic framework from which it is possible to derive hypotheses concerning the performance of humans. Although some work has been done with adults (8), there is no general agreement at to whether increased motivational factors "help" or "hinder" a specific response. Hypotheses 3 and 4, then, were conceived purely as matters for empirical investigation.

METHOD

Materials

A series of "stimulus cards"—20 simple geometric designs drawn with a rule in black india ink on 4 by 5 in. index cards, a series of "comparison cards"—the same 20 designs drawn freehand in pencil on 4 by 5 in. yellow index cards, and a box of 15 red poker chips constituted the materials utilized in this experiment. The figures on the yellow cards were actually drawn by the experimenter, but were presented to each subject as the work of a "fourth (or fifth) grader just like yourself." The reason for this procedure is discussed below.

Subjects

The subjects used in this investigation were 120 white fourth and fifth grade children, 60 boys and 60 girls, with a mean CA of 10 years, 0 months, a mean IQ of 110.4, and a mean grade equivalent of 4.5. They were all obtained from one elementary school in Cedar Rapids, Iowa. Only "normal" children were included in the study; i.e., children who were, according to the principle's report, not undergoing any kind of psychological treatment procedure, were not crippled, etc. A superficial analysis of fathers' occupations revealed that the children came from predominantly middle-class homes.

Experimental Design

The subjects were assigned randomly to three major reinforcement groupings, and again randomized into motivational subgroups within each of the main groups. Group I received continuous positive reinforcement for trials 1 to 10, followed by a series of 10 nonreinforced trials; Group II received positive reinforcement only on trials 1, 4, 5, 7, and 10, followed by a series of 10 nonreinforced trials; and Group III received a continuous series of 20 nonreinforced trials.

Procedure

Each subject was seen individually in a private office and was read the following instructions:

> I'd like to play a game with you today. Other boys and girls have played this game with me and have liked it a lot. I'd like to see how you do at it.
>
> On this sheet of paper I'd like you to draw a design which I shall show you in just a moment. You can draw it anywhere at all on the page and you

> can make it any size you want. Many other children in the fourth (fifth) grade have drawn these designs for me, and each time I present a new design to you on a white card like this (demonstrated with blank card), I'll show you on a yellow card like this (demonstrated) what kind of a job a fourth (fifth) grader just like yourself can do. Now, if you do a better job than this (pointing to yellow comparison card) I'll put a big plus mark on your paper like this (demonstrated), and if you do a poorer job I'll put a minus like this (demonstrated). And if you do just about as well I won't put anything on your paper. Get it? A plus mark for a better job, a minus mark for a poorer job, and nothing at all if you do just about as well. Draw these designs as quickly as you can, but work carefully since you won't be allowed to erase.
>
> Before you begin, though, I'd like you to tell me what you honestly think you'll get for this first drawing, and to make it more fun for you I want you to use these chips to tell me. (Box and chips moved to center of table in front of subject.) Here is a pile of 15 chips which you are to use to tell me how certain you are of getting a *plus* mark on your paper. You can bet with as many or as few chips as you like to tell me how certain you are of getting a *plus*. For example, you can bet up to 15 of the chips if you're absolutely 100 per cent certain you're going to get a plus; or you can bet just *one* of the chips if you're not certain at all of getting it. Or you might want to bet somewhere in between, like 5, 6, 7, 8, or 9. Remember, the more certain you are of getting a plus on your paper the more chips you will want to bet. The less certain you are of a plus the fewer chips you'll want to bet. You can use from none at all to all 15 of them. You will make a bet every time before you draw. Count out your chips carefully, and put them in this box. Any questions? [3]

If the subject was part of a low motivation group he was then asked: "How certain are you of getting a plus mark on this first design?" If he was a high motivation subject he was read the following set of motivating (i.e., "ego-involving") instructions:

> Now I'm interested in knowing who shall do better in your class, the boys or the girls. I know you'd like to see the boys (girls) come out on top, wouldn't you? All right, then, I want you to try to do the *very best* job you can. Also, I'm interested in knowing which fourth (or fifth) grade is best in Cedar Rapids. I'm sure you'd like your class to come out on top, so remember—do your very best! Now, how certain are you of getting a plus mark on this first design?

The subject then indicated the certainty of his prediction of receiving a plus mark for the first performance by counting out into the box whatever amount of poker chips he desired, after which a sheet of paper and the first pair of stimulus and comparison cards were presented. The comparison card was removed from the subject's vision after about a 4-second exposure, and then placed where only the experimenter could see it. The purpose of introducing such a card in the first place was to provide some "face validity" to the drawing game in general, and to prevent the subject from feeling the experimenter's evaluations were merely arbitrary.

---

[3] This betting method is a modification of one used by Castaneda (1). With this technique direct verbalization of prediction is avoided. Providing the child with chips that must be manipulated requires him to do some "subjective scaling" and careful thinking regarding each bet he makes. Nothing is won or lost with these chips. They were introduced solely as a technique of slowing down the subject so as to make him think more carefully about his response.

After the subject had copied the stimulus figure the experimenter would pick up the paper, glance at it carefully, look over with a definite head movement to where the comparison card lay, glance back again at the subject's drawing and place either a large red plus mark on the paper or do nothing at all. A fresh sheet would then be put in front of the subject and the entire procedure begun again. While the child was busy copying the stimulus card, the experimenter would remove the chips counted out into the box and replace them in a single symmetrical pile with the rest of the chips on the table. After completion of the 20 trials each child was told he had done, on the whole, a very excellent job.

RESULTS

The criterion measures used for testing the hypotheses of this study were the numbers of poker chips bet by a subject to indicate the degree of certainty with which he made a prediction of success for himself on each of the 20 trials constituting the experimental session.

Figure 1 represents schematically the results obtained on the three experimental groups. Each point on the curve is based on the mean number of chips bet (mean CPR's) by all 40 subjects in one group. An inspection of the curves at trial 1 reveal that mean CPR's do not differ significantly at that point, but that differences do manifest themselves with the advent of subsequent trials.

Trials 1 to 10

Using the technique of trend analysis for the data obtained for trials 1 to 10 the following results were obtained: (1) *Fs* of 25.15 for the 100% group, 3.93 for the 50% group, and 5.01 for the 0% group (all significant at less than the .001 level of confidence) revealed there were highly significant increments in mean CPR's from trials 1 to 10 under the three reinforcement schedules. (2) Tests for differences between pairs of curves revealed that significantly greater increments in mean CPR's were demonstrated under the 100% in comparison to the 50% or 0% schedules (significant at less than the .001 level of confidence), with no significant difference appearing between the 50% and 0% schedules. (3) No statistically significant differences appeared between the high and low motivation groupings.

Trials 11 to 20

It will be remembered that 10 successive nonreinforcements were presented to all subjects after they had been administered either 100%, 50%,

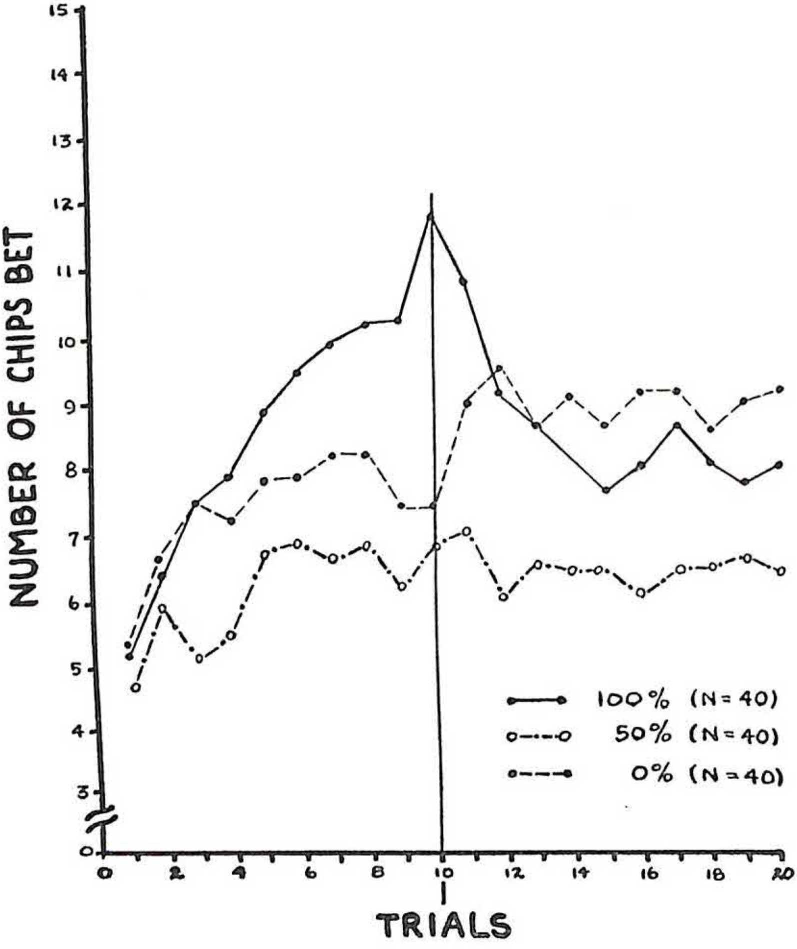


Figure 1—Mean certainty of prediction responses for all groups.

or 0% reinforcement for trials 1 to 10. Analysis of variance of the data revealed that the trials by reinforcement interaction (a test of whether the three groups manifested differential decrements over trials 11 to 20) barely missed statistical significance at the .05 level of confidence. This may have been due to the fact that the 50% and 0% groups, though evidencing no extinction trends in and of themselves, were contributing additional sources of variance to the total extinction data. Because of this insignificant T × R interaction it was not deemed feasible to test for differences between pairs of curves as had been done with the data obtained for trials 1 to 10. However, since an inspection of the empirical curve for the 100% group revealed a decided decline in mean CPR's from trials 11 to 20 whereas the 50% curve did not, it was decided to test for the presence of a 100% extinction curve. A sigma of the difference between the means of trials 11 and 20 netted a critical ratio of 3.48, which is significant at the .002 level of confidence. Thus, it appears a significant decrement does manifest itself for the 100% group, but this finding must be viewed with caution in light of the insignificant T × R interaction.

As with data for trials 1 to 10, comparisons of motivation results for trials 11 to 20 proved statistically insignificant.

DISCUSSION

Results, then, supported that part of hypothesis 1 which posited that a group receiving 100% reinforcement would evidence greater increments in mean CPR's during trials 1 to 10 than would a group receiving either 50% or 0% reinforcement. This result conforms to those obtained from more conventional partial reinforcement studies using adult humans as subjects (2, 4, 5), which have demonstrated that acquisition of a response is directly proportional to the frequency of reinforced trials, proceeding, generally speaking, somewhat more rapidly under continuous than partial reinforcement.

On the basis of the findings of such studies it was further hypothesized that the group receiving 0% reinforcement during the first 10 trials would evidence statistically weaker increments in mean CPR's than would be the case with either the 50% or 100% groups. That this is not true is evident from an inspection of the curves in Figure 1. Not only was a significant increment trend demonstrated, but also there was no reliable difference between it and that of the 50% group.

Post hoc considerations of this somewhat paradoxical result lead to two tentative hypotheses, the first being that a state of anxiety was produced in these 0% subjects by the continuing series of nonreinforced trials (which may have been construed by the subjects as failure). If (*a*) anxiety is thought of as a drive, (*b*) it is considered that increased drive interacts with existing tendencies-to-respond so as to increase probability of strong response, and (*c*) the CPR can be thought of as a response of this nature, then the results for the 0% group can perhaps be considered from the point of view of such current psychological formulations as the Spence-Taylor type (9). Second, it might be that the increased CPR was due to some anxiety-reducing mechanism, perhaps of the "denial" variety. That is, it is possible these subjects tended to reject the reality aspects of the situation in which they were involved, steadily increasing their CPR's in spite of continued nonreinforcement. Because of a lack of opportunity to contrast positive and nonreinforcement, as was possible for the 50% group, the entire experimental situation was highly ambiguous for 0% subjects. Hence, for them an anxiety-reducing mechanism such as denial would be more "available" than for the 50% group subjects, who have experienced both positive and nonreinforcement, and who would not be as likely to respond negatively to the reality aspects of the situation.

Results offered no statistical support to hypothesis 2 which predicted significantly greater decrements in mean CPR's during trials 11 to 20 for the 100% group than for either the 50% or 0% groups. There is a suggestion, however, from an inspection of the 100% and 50% curves and a consideration of the significant extinction test for the 100% data that results of these two groups are not completely deviant from the general conclusion reached in more conventional partial reinforcement research on human and animal subjects, namely, that a response will be more resistant to extinction if acquired under a schedule of partial rather than continuous reinforcement.

As was stated above, there was no systematic framework from which it was possible to predict the performance of high vs. low motivation groupings. The two motivation hypotheses in this study, therefore, were almost purely speculative, and were devised solely for purposes of empirical investigation. It will be remembered that an attempt was made to induce a high state of motivation in certain subjects by presenting a set of what were considered "ego-involving" instructions. It seems likely these instructions did not serve the purpose for which they were intended. Instructing a child at the beginning of the experimental situation to do his very best so that his sex will win top position in his class and his class top position among other similar classes in the city's schools apparently does not induce or sustain the desired state of motivation or drive. Were this experiment replicated it would perhaps be better first to match high and low anxiety groups on the basis of a technique such as the Iowa Anxiety Scale for Children and then induce motivation with a set of instructions geared to be more highly ego-involving. Such instructions could perhaps tell the child that the results of his performance during the experiment would be reported to his teacher, who would then incorporate them in a final grade for him.

To summarize, what has been demonstrated in this experiment is that Child A, instructed that three physical events are possible of occurrence during the course of the experimental situation (receipt of a plus, a minus, or nothing at all for a performance), will, significantly, become increasingly more confident that future success will occur as he receives a series of continuing successful experiences than will Child B, who is administered a number of success experiences interspersed with an equal number of nonsuccess experiences (defined as the receipt of neither a plus nor a minus), or Child C, who receives a series of continuous nonreinforcements. This is interpreted as a function of the greater frequency of successes experienced by Child A. Though not statistically significant, there is a tendency for certainty to diminish somewhat more rapidly for Child A than for B or C as each is administered the same number of continuous nonreinforcements after the first series of trials, with the latter two children manifesting no significant decrements in response whatsoever. Child C, who does not experience success at all during the entire course of the experiment will, nonetheless, evidence steady increments in the certainty with which he predicts the occurrence of potential success for himself from trial 1 to 10. This is interpreted as a function of the presence of anxiety factors which are operating in this child.

SUMMARY

This research was devised to investigate the acquisition and extinction of a response, namely, the degree of certainty with which a child makes a prediction concerning the occurrence of a success experience for himself, under three schedules of reinforcement (100%, 50%, and 0%) and two levels of motivation (high and low).

One hundred twenty normal fourth and fifth grade children were assigned randomly to three major reinforcement groups, and again divided into six motivational subgroups. Group I (100%) was administered a series of continuous positive reinforcements during trials 1 to 10, followed by a series of 10 nonreinforced trials; Group II (50%) received positive reinforcements on only five randomly selected trials during acquisition, followed by 10 nonreinforced trials; and Group III (0%) received a continuous series of 20 nonreinforced trials.

Each subject, seen individually, was presented with a "game" involving the copying of a series of 20 geometric designs. Before starting, he was informed that one of three possible events would occur after the completion of each individual drawing, the receipt of a plus, a minus, or nothing at all. Before copying each design he was instructed to "make a bet" from a pile of 15 poker chips to indicate how certain he was of the occurrence of a success experience for himself (i.e., the receipt of a plus). Any number of chips from 0 to 15 could be bet, zero indicating a success prediction of least and 15 a success prediction of most certainty. Nothing was won or lost with these chips; they were introduced solely as a nonverbal technique of obtaining a child's certainty of prediction responses (CPR's), and of getting him to do some "subjective scaling" regarding each bet. All subjects in "high motivation" subgroups were provided, additionally, with a set of specially prepared "ego-involving" instructions to induce a state of drive.

The criterion measures used in statistical analyses were the numbers of poker chips bet by each subject to indicate the degree of certainty with which he made a prediction of success for himself on each of the 20 trials constituting the experiment. In general, results yielded by the acquisition data are considered consistent with the findings of continuous and partial reinforcement research (using animals and adult humans as subjects) of somewhat more rapid acquisition of behavior under continuous than partial reinforcement. Though missing statistical significance slightly, results of the

100% and 50% extinction data are in the same direction as the general conclusion reached in these same studies, indicating greater resistance to extinction when behavior has been acquired under a schedule of partial as compared to continuous reinforcement. No significant differences between high and low motivation groups were demonstrated.

References

1. Castaneda, A. A systematic investigation of the concept expectancy as conceived within Rotter's Social Learning Theory of Personality. Unpublished doctor's dissertation, Ohio State Univer., 1953.
2. Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 1951, 42, 1-5.
3. Grosslight, J. H., & Child, I. L. Persistence as a function of previous experience of failure followed by success. *Amer. J. Psychol.*, 1947, 60, 378-387.
4. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, 25, 294-301.
5. Humphreys, L. G. The effects of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid responses. *J. exp. Psychol.*, 1939, 25, 141-158.
6. Jenkins, W. O., & Stanley, J. C. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, 47, 193-234.
7. McCandless, B. R., & Rosenblum, S. Psychological theory as a determiner of experimental pattern in child study. *Rev. educ. Res.*, 1952, 25, 496-525.
8. Mosteller, F., & Nogee, P. Measurement of utility in a gambling situation. Reported in Jenkins and Stanley (6).
9. Spence, K. W., & Taylor, Janet A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, 42, 183-188.

AN OBJECTIVE METHOD FOR THE MEASUREMENT AND ANALYSIS OF CHILD-ADULT INTERACTION[1]

CLARK E. MOUSTAKAS, IRVING E. SIGEL, and HENRY D. SCHALOCK[2]
The Merrill-Palmer School

Interaction behavior is an important indicator of the underlying relationship between individuals. On the basis of overt behavior we often make inferences and interpretations and arrive at an understanding of its meaning and significance for the individuals involved. This kind of knowledge is possible when we have accurate presentations and descriptions of observable behavior.

How are we to collect the material from which more basic and dynamic understanding may be developed? What are the facts of interactional behavior which we can record, classify, and examine, and from which we may establish testable hypotheses? Observation appears to be the most appropriate technique for the descriptive phase of such a study. A method is needed by which reliable descriptive material on adult-child interaction can be obtained.

Our purpose here is to present a procedure for objective description and recording of adult-child interaction. Data are presented to show the usefulness of the procedure under three types of interpersonal situations. This

---

[1] The authors express their appreciation to The Merrill-Palmer School and its Director, Dr. Pauline Park Wilson Knapp, for making the research possible; to the parents who cooperated in the project and gave generously of their time and support; to Dorothy Haupt, Leader, Preschool Groups Service, who facilitated our approach to the children and assisted with the scheduling; to Drs. Chester D. Clapp and Esther B. Frankel, who introduced the parents to the observation procedures and laboratory; to Joan Godshalk, graduate student, who served as an observer and helped to establish reliability of the schedule; and to Dorothy L. Tyler, Editor of Publications, who prepared the manuscript for publication.

[2] Drs. Moustakas and Sigel are psychologists on the staff of the Merrill-Palmer School; Mr. Schalock, now on the faculty, Oklahoma A. & M. College, was a graduate student, 1953-54.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

procedure involves the use of a predetermined set of general purpose observational categories which can be applied to different types of adult-child interaction.

A major source of suggestions for these categories was the excellent study of Barbara Merrill Bishop (2). Her schedule for studying mother-child social interaction was used as a point of departure. We then made refinements, clarified category definitions, and added numerous categories to make our schedule more inclusive, detailed, and widely applicable. The additional categories were in part selected from other studies using a system of behavioral categories (3, 4, 5, 6, 7, 8) and in part were based on the authors' experience. A scale for recording anxiety and hostility was added. Further revisions, suggested by experience in using the schedule to observe mother-child and therapist-child interaction, are incorporated in the final schedule.

SELECTION OF CATEGORIES

Selection is extremely important in establishing a set of categories for observation. It is often based on the nature of the problem and on a specific theoretical framework. Though the present schedule was constructed on an empirical basis, in order to develop a schedule that would have wide application, certain theoretical assumptions are nevertheless inherent in the schedule. The primary assumption is that interactional behavior is manifold and composed of various types of observable behaviors. Another assumption is that it is possible to construct overt behavior categories that will not only have wide applicability for various kinds of adult-child interaction but will also be of use and value to those holding different theoretical points of view.

The criteria used in selecting the categories in the schedule were: comprehensiveness, relevance and meaningfulness, and ease of identification.

Comprehensiveness. This criterion refers to the categories *in toto,* that is, the question of whether they embrace a relatively inclusive number of adult-child interactional behaviors. Review of the literature, together with the experience of observers and writers, was the basis for ascertaining what categories should be selected to cover the major aspects of such interactional behavior.

Relevance and meaningfulness. Certain arbitrary decisions were made as to the contribution certain categories made in furthering our understanding of adult-child interaction processes. Here our own theoretical biases concerning adult-child roles, expectations, and consequent behaviors influenced our decisions and are evident in the specific categories selected.

When an adult interacts with a child he may influence the child in many different ways to bring about the kind of behavior he wishes. He uses the type of behavior appropriate to his goal. An adult has a certain status position derived from his maturity, size, and function. As a mature, socialized individual he can use his position to frustrate the child's needs or to gratify them. He can initiate certain kinds of behavior in the child and ignore others. Certain categories were derived on the basis of these influence characteristics of the adult.

Although a child has relatively limited size, maturity, and socialization, he can also influence the adult. He can frustrate the adult, intentionally or otherwise; he can also direct and gratify the adult. Thus, though not powerful, not mature, and not operating with the same general wherewithal, the child can influence adult behavior. These child influence expressions are also a source of categories in the schedule.

In order to keep the schedule of manageable length, we included a category only when it offered information that would help us more adequately to understand our own conception of adult-child interaction.

Ease of identification. This criterion can be applied only when categories can be precisely, clearly, and adequately explained. That is, trained observers must find the relevant behaviors relatively easy to identify and label. For this reason certain details of behavior were omitted from the schedule. The greater the ease with which observers could label the observed behavior, the greater the chance for observer agreement.

The Schedule: Discussion and Analysis

Schedule categories for observing and scoring adult behavior and child behavior are mainly the same.[3] Definition, illustration, and discussion as applied to child and adult are therefore presented together in this section under the various category headings. Table 1 presents a complete list of categories, with scoring symbols.

Attention Categories

Nonattention. *A* (adult) directs his attention to something other than *C* (child), or *C* directs his attention to something other than *A*.[4] Ex: *A* reads a book or leaves the room for a personal reason. Ex: *C* plays quietly with back to *A* or stares out the window. Nonattention may be accompanied by humming, singing, or talking, but this behavior is not apparently directed toward *C* or *A*. No interactive characteristics are observable.

---

[3] Only two categories differ: category *z*, Interpretation, broadly defined in the child categories and defined in greater detail and more specifically in the adult categories; *x*, Threat of Attack, is omitted from the adult categories.

[4] After reliability was established, a broad rating scale was constructed to provide a general measure of the frequency of negative and positive expressions in the noninteractive behavior of *A* and *C*. The rating is based on tonal quality, choice of gestures, words, and the content of the expression, as follows: 0 Behavioral expression appears neither positive nor negative. 1 Behavioral expression is positive, e.g., "What a lovely doll!" (*C* caresses the doll.) 2 Behavioral expression is negative, e.g., hitting or kicking objects, tenseness of facial muscles while drawing, etc.

Attentive Observation. *A* noticeably directs his attention to *C* and/or *C*'s activity by silently watching; or *C* directs his attention to *A* in this way. Ex: *A* watches *C* as *C* plays in the sand. Ex: *C* watches *A* as *A* draws a picture.

Recognition. *A* responds to *C*'s stimulation in a way that indicates definite awareness of *C*'s statement or activity; or *C* responds to *A*'s stimulation in this way. Ex: Mm hm. Ex: Yes, I see. Ex: (*A* listens and generally attends to *C* as *C* tells about an experience.)

These two categories, Attentive Observation and Recognition, convey interest and concern on the part of adult or child. Attentive Observation may have an important influence on the behavior of child or adult, but this influence is not as clear as that associated with Recognition. While *A* is attentive to *C*, *C* may be unaware of *A*; but when *A* recognizes *C*, this recognition must be based on a stimulus directed to *A*. As defined in this schedule, Recognition is always interactive, while Attentive Observation may be unidirectional.

Statement of Condition or Action. A comment describing an existent state, situation, or action. Ex: It's hot in here. Ex: I'm afraid. Ex: I'm going now. Though such statements of condition or action may have stimulus properties, from the observer's point of view they apparently do not influence behavior or elicit particular responses.[5]

Joint Participation in Activity. *A* and *C* are mutually engaged in an activity. Ex: *A* and *C* are reading, working with puzzles, singing, playing, drawing, or painting together.

Offering Information. *A* or *C* offers knowledge or guidance. (1) Information is offered verbally. Ex: *C*: Why won't this open? (Referring to cash register.) *A*: You have to press the keys hard to make it open. Ex: *C*: You have to unsnap my boots before you can take them off. (2) Information is offered by means of demonstration. Ex: (*C* is apparently trying to make wheels with tinker toys and is having difficulty.) *A*: Let me show you how to do one. (*A* proceeds to demonstrate how to make a wheel.) Ex: *C*: Let me show you how to work this puzzle. (*C* demonstrates.) (3) Information is offered through the combined methods of demonstration and explanation. Ex: (*C* is having difficulty working puppets. *A* takes puppet, shows *C* how to work it.) *A*: You put your thumb in here and work it like this. Ex: (*A* asks *C* how the rifle works.) *C*: Here's how you shoot it. You keep pumping it until the ball flies out.

Giving Help.[6] Physical aid is given by *A* or *C*. Ex: (*C* is trying to hook the cars of a train together but is having no success. *A* holds cars while *C* hooks them together.) Ex: (*C* holds finger on string while *A* ties a knot.)

---

[5] After reliability was established, and on the basis of further experience, the category Statement of Condition or Action was subdivided as follows: b_1 or b_1 Statement involving a situation external to the individual. Ex: I went to the circus yesterday. Ex: It's cold outside. b_2 or b_2 Statement involving an emotional state of the person. Ex: I'm afraid. Ex: I hate my father. b_3 or b_3 Statement of action or proposed action. Ex: I'm making a scrapbook. Ex: I'm leaving now.

[6] After reliability was established, and on the basis of further experience, the category Giving Help was subdivided as follows: h_1 or h_1 *A* helps *C* in completing a task; or *C* responds to *A* in this way. Ex: *C* helps *A* put tinker toys away. Ex: *A* holds finger on string while *C* ties a knot. h_2 or h_2 *A* helps *C* by completing a task for *C*; or *C* responds to *A* in this way. Ex: *A* makes tinker toy objects for *C*.

These two categories, Offering Information and Giving Help, both involve assistance to the child or adult. However, a distinction may be made. Information involves a kind of teaching, through demonstration, explanation, or the stating of a principle or fact. Helping, as used here, means actually doing or helping to do the task at hand.

Reassurance. Comfort or encouragement is offered by *A* or *C*. Ex: *A*: Don't feel too bad. Most children find it difficult to work that puzzle the first time. Ex: *A*: It doesn't matter that you spilled some paint on the floor. It can be easily washed. Ex: (*A* drops a piece of doll furniture, which breaks.) *C*: Don't worry, I've got some other furniture to play with.

This category, Reassurance, differs from the categories Giving Help and Offering Information in not involving information or assistance relevant to solving a problem. It is a method of offering assistance through emotional support.

Stimulus Categories

Several categories involve an attempt to elicit a particular kind of response or behavior by *A* or *C*. These stimulus categories are discussed next.

Seeking Information. *A* or *C* asks for information. (1) Personal information is sought, i.e., information about parents, sibling or peer relationships, feelings toward school, etc. Ex: *A*: How is everything in school these days? Ex: *A*: Does your little brother break your toys? Ex: *C*: Mommy, are you mad at Daddy? (2) Information is sought concerning the ongoing activity. Ex: *A*: Do you like that puppet more than the others? Ex: *C*: What is this thing?

Seeking Help.[7] *A* or *C* asks for help. (1) *A* seeks assistance in completing a task; or *C* responds to *A* in this way. Ex: Help me carry these chairs. (2) *A* requests *C* to complete a task for him; or *C* responds to *A* in this way. Ex: *C*: Will you pound this in for me? I can't do it right.

Seeking Reassurance. Comfort or encouragement is sought by *A* or *C*. Ex: Are you sure you're not angry with me? Ex: Am I doing this the way I should?

Seeking Recognition. Attention to activities, productions, or statements is sought by *A* or *C*. Ex: See what I've made? Ex: Look what I'm doing.

Seeking Praise. Desire for praise is indicated by *A* or *C*. Ex: Don't you think this is a nice picture? Ex: How do you like the way I fixed your hair?

Seeking Affection. Expressions of love or tenderness are sought by *A* or *C*. Ex: Come and kiss me. Ex: Hold me on your lap.

Seeking Reward. Special privileges or benefits are sought by *A* or *C*. Ex: I've been a good boy all day. Will you give me some candy? Ex: If you let me go downtown, I'll bring you back an ice cream cone.

Seeking Permission. Consent is sought by *A* or *C* for a proposed activity. Ex: Is it all right if I leave for a minute? Ex: Could I play with all these dolls?

Orienting and Directing Categories

Another group of related categories in the schedule are Orienting, Directing, Restricting, and Forbidding. Orienting sets the stage and usually permits a wide range of behavior. Directing tends to focus attention or activity on specific behavior or in a particular direction, and is narrowly oriented. Restricting and Forbidding are also focused on particular behavior in a specific way, narrowly limiting or completely stopping the behavior.

---

[7] Revised form given. The earlier form read: *A* or *C* requests physical assistance. This may be done either verbally or nonverbally. Ex: Help me carry these chairs. Ex: I can't pound this right. Will you hold this peg for me?

Orienting. Limits, boundaries, and roles are indicated by *A* or *C*.[8] (1) Boundaries of the situation are indicated by *A* or *C*. Ex: You may do whatever you like here. (2) *A* or *C* structures time. Ex: You have only about five more minutes left to play. Ex: You can stay only a few more minutes. (3) Roles are structured by *A* or *C*. (*a*) Role is indicated on the basis of status or power by *A* or *C*. Ex: C: How do you know I'll get sick if I eat candy? *A*: Older people always know these things. Ex: *C*: Mommy, you be the sheriff and I'll be the cowboy. Ex: *A*: You'll have to ask daddy. He makes the rules. Ex: *C*: I'm the smallest, so I get my turn first. (4) Roles are indicated by leaving responsibilities to *C* or *A*. Ex: I can't tell you what to do, you must decide for yourself. Ex: How do you think it should be done? Ex: You can use these things any way you want.

Directing. *A* attempts to influence *C* by command or suggestion; or *C* attempts to influence *A* in this way. (1) *A* or *C* uses verbal or nonverbal suggestion. Ex: You could play with the darts if you wanted to. Ex: Would you like to sit over here? Ex: (*C* is hitting table top with hammer. *A* places pounding block on table.) (2) *A* or *C* gives a command accompanied by a threat. Ex: Clean up this mess or you can't play in here any more. Ex: If you don't bring me that knife, I'll throw this at you. (3) *A* or *C* gives a command. Ex: Put the gun on the table. Ex: Get me that book.

Restricting. *A* attempts to modify *C*'s behavior by reducing the intensity, speed, manner of executing, etc., but apparently does not intend to stop the activity completely; or *C* attempts to modify *A*'s behavior in this way. (1) *A* attempts to restrict *C*'s ongoing activity by warning; or *C* responds to *A* in this way. Ex: Be careful, or you will fall and hurt yourself. Ex: You hadn't better come any closer, or I might splash paint on you. (2) *A* attempts to restrict *C*'s ongoing activity, while offering an explanation; or *C* responds to *A* in this manner. Ex: Don't splash the water so high, Jimmy, or you'll get your clothes wet. Ex: Don't open the window so high, or I'll get cold. (3) *A* attempts to restrict *C*'s ongoing activity by threat of bodily punishment; or *C* threatens *A* in this way. Ex: It's all right to paint, but if you spill any more on the table you'll get spanked. Ex: Don't talk so loud, or I'll slap you. (4) *A* attempts to restrict *C*'s ongoing activity by threat of loss of affection; or *C* attempts to restrict *A* in this way. Ex: I can't be nice to you when you yell so loud. Ex: I won't sit on your lap if you don't read slower. (5) Restriction of the ongoing activity is attempted by *A* or *C* through a direct statement. Ex: I don't think you should pound the table quite so hard. Ex: Don't make me go so fast. (6) *A* recognizes or clarifies *C*'s feelings while attempting to restrict *C*'s ongoing activity; or *C* responds to *A* in this way. Ex: I know you'd like to eat all those cookies, but you may only have two. Ex: I know you're very mad at me, Mommy, but don't push me so hard. (7) Restricting the ongoing activity by limiting the use of objects is attempted by *A* or *C*. Ex: You may fill that truck only once more.

---

[8] After reliability was established, and on the basis of further experience, all subcategories of Orienting were included in the observation schedule for the child.

Ex: You may use only four blocks to build the house. (8) Restricting the ongoing activity by physical restraint is attempted by *A* or *C*. Ex: (*C* attacks by hitting. *A* attempts to regulate the force of the blow.) Ex: (*A* slaps *C*. *C* breaks the force of the blow.)

Forbidding. An ongoing activity is interfered with by *A* or *C* with the apparent intent of stopping it completely. (1) *A* attempts to stop *C*'s ongoing activity while offering an explanation; or *C* attempts to stop *A* in this way. Ex: I can't let you paint on the walls, because the paint won't come off. Ex: Stop putting my toys away. I want to play with them. (2) *A* attempts to stop *C*'s ongoing activity while offering an alternative; or *C* responds to *A* in this manner. Ex: I can't let you shoot me in the face, but you can shoot me in the body if you want to. Ex: You can't have that chair, but you can sit on the other one. (3) Feelings are recognized and clarified by *A* or *C* while an attempt is made to stop an ongoing activity. Ex: I know you feel like hurting me, but I can't let you do that again. Ex: I know you don't like me telling you what to do, but stop saying what I say. (4) *A* or *C* attempts to stop an ongoing activity by threat of bodily punishment. Ex: If you don't stop making so much noise, I'm going to spank you. Ex: You shut up, or I'm going to hit you. (5) *A* attempts to stop *C*'s ongoing activity by threat of loss of affection; or *C* responds to *A* in this manner. Ex: I won't like you any more if you don't stop splashing me with water. (6) *A* or *C* attempts to stop an ongoing activity by means of a directly phrased statement. Ex: I will not permit you to hit me with that again. Ex: Stop that! (7) *A* attempts to stop *C*'s ongoing activity by proceeding to remove objects from *C*; or *C* responds to *A* in this manner. Ex: *C* is pounding on window with toy hammer. *A* takes hammer from *C*. Ex: *A* is writing on a note pad; *C* takes pencil from *A*. (8) *A* or *C* attempts to stop an ongoing activity by employing physical barriers. Ex: (*A* closes door when *C* tries to leave room.) Ex: (*C* blocks *A*'s path when *A* attempts to get coat and hat.) (9) *A* or *C* uses physical restraint in attempting to stop an ongoing activity. Ex: (*C* slaps at *A*. *A* hold's *C*'s arms, stopping the attack.) Ex: (*A* starts to leave room. *C* stops *A* by pulling on coat.) (10) *A* attempts to stop *C*'s ongoing activity by warning; or *C* attempts to stop *A* in this way. Ex: If you don't stop swinging those scissors, you'll hurt yourself. Ex: If you don't stop reading, you'll get a bad headache.

Criticism, Discipline, Rejection Categories

The next five categories in the schedule are Criticism, Disciplinary Action, Physical Attack, Threat of Attack, and Rejection. Threat of Attack, when it occurs without apparent provocation, is used only in observation of the child's behavior. Threats of attack with apparent provocation are scored in the specific categories involved; i.e., restricting or forbidding.

Criticism. *A* gives a negative evaluation of *C* or *C*'s productions; or *C* responds to *A* in this way. (1) *A* makes a negative evaluation of *C* by employing shame, ridicule, or disparaging remarks; or *C* responds to *A* in this manner. Ex: Don't be so stupid. Ex: Why don't you act your age? Ex: Damn you! Ex: Ha! Ha! Spilled milk all over yourself. (2) *A* makes a negative evaluation of *C*'s productions by employing shame, ridicule, or disparaging remarks; or *C* makes negative evaluation of *A*'s productions in this way. Ex: That's a sloppy way to paint. Ex: Aw, this isn't any good! You don't know how to make pictures like my teacher does. (3) *A* makes a negative evaluation of *C* and *C*'s productions by employing shame, ridicule, or disparaging remarks; or *C* responds to *A* in this way. Ex: When you talk like that, you are acting like a baby.

As used in the schedule, Criticism always involves some kind of negative evaluation; therefore, constructive criticism is not included in this category. When constructive criticism occurs in a child-adult interaction, it may involve giving information, suggesting, directing, etc., and is scored in the specific category involved.

Disciplinary Action. *A* attempts to inflict pain on *C* as a response to an indicated undesirable act of *C*; or *C* responds to *A* in this way. The concept of punishment here usually includes moral, ethical, or teaching values, but it may involve solely the intent to inflict pain or hurt. (1) *A* punishes *C* by isolation; or *C* punishes *A* in this way. Ex: I've told you not to do that before. Now you go to your room and stay there. (2) *A* punishes *C* by deprivation of physical objects, food, privileges, or restriction of space; or *C* responds to *A* in this way. Ex: Just for that, you can't play with the gun any more. Ex: I told you not to read. (*C* takes paper from *A* and hides it.) (3) *A* punishes *C* by destroying or attempting to break objects in response to *C*'s apparently irritating behavior; or *C* responds to *A* in this way. Ex: (*C* breaks toy after seeing *A* kiss baby.) (4) *A* punishes *C* by means of corporal punishment; or *C* punishes *A* in this manner. Ex: (*C* sasses *A*.) *A*: That does it. (*A* slaps *C*.) Ex: Just for that I'm going to spank you. (*C* proceeds to spank *A*.) (5) *A* punishes *C* by means of deprivation of love or affection; or *C* punishes *A* in this way. Ex: Since you won't sit still, I'm not going to hold you. Ex: Since you won't let me go, I'm not kissing you good-by.

Physical Attack. *A* strikes *C*, or *C* hits *A* without apparent provocation. Ex: *C* kicks *A*.

Threat of Attack. *C* threatens to attack *A* without apparent provocation. (1) *C* verbally threatens to attack *A*. Ex: I'm going to hit you. (2) *C* threatens to attack *A* by gesture. Ex: (*C* approaches *A*, holding a toy hammer in striking position.) (3) *C* both verbally and gesturally threatens *A*. Ex: I'm going to shoot you with this rifle. (Aims air rifle at *A* and moves toward *A*.)

Rejection. *A* indicates nonacceptance of *A* and/or *C*'s stimulation, either verbally or gesturally; or *C* responds to *A* in this way. (1) *A* changes the subject or interrupts *C* with irrelevant conversation; or *C* responds to *A* in this manner. Ex: (*C* asks *A* to help him build a swing with tinker toys. *A* changes the subject, saying, "You haven't painted today yet.") Ex: *A*: Pick up the blocks. Ex: *C*: It's cold in here, isn't it? (2) *A* denies the validity of *C*'s statement or action; or *C* responds to *A* in this manner. Ex: *C*: I'm a good boy. *A*: No, you are a bad boy. Ex: *A*: That's a house. *C*: No, it's not a house. Ex: (*A* shakes head negatively.) (3) *A* rejects *C* as a person; or *C* rejects *A* as a person. Ex: I don't want anything to do with you. Get away from me. Ex: I hate you. Ex: (*C* asks *A* to play with him. *A* walks away.) (4) *A* makes a direct appeal to *C*, and *C* ignores or evades him; or the appeal is made by *C* and *A* ignores or evades him.[9]

---

[9] This category was originally classified as Nonrecognition, and later included in the group of rejection categories. Nonrecognition is coded as n in Table 2.

Ex: (*C* looks at *A* and asks, "Do you like me?" *A* does not respond.) Ex: *A*: Don't you think this is a nice picture I have made for you? (*C* does not respond.)

Approval or Reward Categories

Four approval or reward categories are included in the schedule. They are Permission, Praise, Affection, and Reward.

Permission. *A* consents to *C*'s proposed activity; or *C* consents to *A*'s proposed activity. (1) Permission is granted with conditions. Ex: Yes, you may use the paints if you put on an apron. Ex: Yes, you may go if you come back in a few minutes. (2) Permission is granted without condition. Ex: Yes, you may have another sheet of paper. Ex: Yes, you may leave now.

Praise. *A* expresses approval of *C* and/or *C*'s productions; or *C* responds to *A* in this way. Ex: That's a fine picture you've made. Ex: You can sure write well.

Affection. *A* gives overt expressions of love or tenderness to *C*; or *C* responds to *A* in this manner. (1) Affection is expressed by physical means. Ex: (*A* rubs or pats *C*'s head.) Ex: (*C* puts arms around *A*.) (2) Affection is expressed verbally. Ex: I love you very much. Ex: I like you. (3) Affection is expressed by physical and verbal means. Ex: *A* hugs *C* and says, "Mm, I like you very much." Ex: *C* kisses *A* and says, "I love you."

Reward. *A* extends special privileges or benefits to *C* as a response to *C*'s behavior; or *C* extends special privileges or benefits to *A*. Ex: For being so good today, you may stay ten minutes longer. Ex: For helping me with this picture, I'm going to give it to you.

Cooperation, Compliance, Noncooperation

These are essentially response categories. Cooperation, which has no coercion aspects, involves deliberate participation, with enthusiasm and willingness. Compliance involves submission and implies coercion and unwillingness. It may or may not be destructive or unreasonable. Noncooperation involves a refusal to participate, an action that goes against the stated request or apparent expectation of child or adult.

Cooperation. *A* responds to *C*'s comments, suggestions, or requests with apparent interest and willingness; or *C* responds to *A* in this manner. (1) Casual or straightforward participation. Ex: O.K., I'll play darts with you. Ex: O.K., I'll put the toys away. (2) Enthusiastic participation. Ex: I would love to play house with you. What could I do? Ex: O.K., that's a very good idea. I'll do it now.

Compliance. Responds to commands, suggestions, or requests with apparent reluctance or indifference. Ex: Oh, all right, I'll do it, but I don't want to. Ex: All right, but I'd rather play with the gun.

Noncooperation. Refusal to accept commands, suggestions, or requests. (1) Refusal to participate accompanied by an explanation. Ex: I won't leave yet. I haven't finished building the house. Ex: I won't play in the sand with you, because I don't want to get my clothes dirty. (2) Refusal to participate accompanied by an alternative. Ex: I won't play in the water with you, but I'll hand you the boats. Ex: I don't want to put the gun on that table, but I'll put it on the ledge. (3) Simple refusal to participate. Ex: No! Ex: I'm sorry, but I won't do that. (4) Strongly expressed refusal to participate. Ex: No! I certainly will not. Ex: No! No! No!

Ambivalent statement. A expresses conflicting ideas, philosophy, or attitudes; or *C* does so. Ex: I don't know whether to let you do that or not. Ex: I like it here, but I'm afraid.

Interpretation

Interpretation, the final category in the schedule, is broadly defined for use in observations of child behavior. Eight separate subcategories are included in the adult schedule. In general, an interpretive response is one which in some respects moves away from expressed statements. As the adult response uses less and less of what the child immediately expresses, the response is more and more removed from the child's awareness and is therefore considered more interpretive. Probably correspondence of phenomenological fields is never identical, and therefore some aspect of interpretation is involved. Individuals respond to communication in terms of their own cognitive organizations, selecting aspects of the stimuli complex as the basis for response.

Interpretation.[10] *A* comments on *C*'s behavior by means of evaluation or analysis, selecting and emphasizing certain aspects of it in terms of theoretical concepts, beliefs, or assumptions, or from knowledge of past history; or *C* responds to *A* in this way. (1) *A* selects an aspect of *C*'s remark and restates the content of the remark. Ex: *C*: That's a car. This is a truck. Here's the baby. *A*: That's the baby, huh? (2) *A* verbalizes the content of *C*'s motor behavior. Ex: (*C* makes four drawings.) *A*: I see that you have made four drawings. (3) *A* restates feelings of verbal statements. Ex: *C*: I don't like that paint. *A*: You just don't like it at all. (4) *A* recognizes feelings in motor behavior. Ex: (*C* pours paint on paper and splashes it around.) *A*: You feel like messing it all up. (5) *A* attempts to clarify *C*'s expressed feelings by elaborating *C*'s statements. Ex: (*C* pours paint from can into container, using more and more paint.) *C*: I love to mix this. *A*: Maybe you feel like putting the whole can in. (6) *A* attempts to clarify *C*'s feelings when the feelings are not immediately verbalized or indicated in *C*'s motor behavior. Ex: (*C* takes red paint and begins to paint mother doll's hair, clutching figure firmly, with intent expression on face.) *A*: Maybe you just feel like smearing her face. (7) *A* attempts to clarify *C*'s feelings, which are not immediately verbalized or indicated in *C*'s motor behavior, by pointing out a temporal relationship between an immediate expression of *C* and some event in the past. Ex: *A*: Yesterday you really knocked the big brother around, but today you seem afraid. Perhaps you would like to hurt him, but are afraid of what he might do to you. Ex: (*C* drinks from nursing bottle.) *A*: You are drinking from the nursing bottle the way you did when you were a baby. It makes you feel good now like you felt when you were a baby. (8) *A* responds to *C* by giving his account of reality to *C*, by explanation and/or translation of symbolic behavior. Ex: *C*: I'm just scared and don't know what it is. *A*: I know. You are afraid because you are alone in this room with a man you do not know and whom you cannot yet trust. Ex: (*C* puts two swings together and points out the inner part, which hangs down and swings.) *C*: Look how it dangles and bumps. (*A* points to dangling swings and says:) Those are two people, your Daddy and Mommy, bumping their thingummies [*C*'s word for genitals] together.

---

[10] The subcategories apply only to observations of adult behavior.

Anxiety-Hostility Ratings

The ratings of anxiety and hostility attempt to measure the intensity of two types of emotional concomitants of behavior. Each rating is based on the style in which the behavior manifests itself. The observer judges the emotional expressions involved by utilizing various clues, such as facial expression, feeling tone of language, body tension, random body movements, tenseness, and tremors. With these manifestations as his bases, the observer judges intensity by ascribing a rating to the specific behavior.

The following ratings are used:

0 = Little or no anxiety or hostility apparent.
+ or − = Indication of some anxiety or hostility.
++ or −− = Indication of much anxiety or hostility.

Such a rating accompanies each of the interactive categories in the schedule.

CONDITIONS AND PROCEDURES FOR OBSERVATION

Laboratory Observations

Laboratory observations were made of mother-child and therapist-child situations in a room set up primarily for play therapy situations, and equipped with a variety of toys, materials, and equipment (9). The observers were seated in an observation booth, wired for sound, behind a one-way vision mirror. The therapist knew that he was being observed, but the children were not aware of the observers. In the mother-child situation, both participants knew they were being observed. The mothers were told that the purpose of the visit was to help us work out methods for observing parents and children in interaction.

The therapist-child situation was based on principles and concepts of child-centered play therapy. The therapist explained to the child that he was free to use the materials in the playroom in any way he wished during the 45-minute session. Mother and child were instructed that they could do whatever they wished while in the playroom, and that they would be notified when the 45-minute observation was over.

Home Observations

Three home observations were made in order to achieve observer agreement and to allow sufficient time for mother and child to become accustomed to the observer's presence (1). Initial practice home observations were made in a family well known to both observers. After the procedures were somewhat familiar, contacts were made with other families. A standardized procedure was employed when a family was to be observed. In a preliminary interview in the home the purpose of the observation was explained and rapport was established with mother and child. The observer also had an opportunity to see the home and so knew in advance where the child usually played, where toys were kept, and what space, if any, was forbidden to the child. An appointment was then made for the initial observation, and the mother was told that during the observation she should proceed with her usual tasks.

Certain home observation principles were established: The observer was to remain close enough to the mother-child interaction to see and hear what occurred, yet as much apart from it as possible. The observers tried to maintain a free, noninteracting relationship with mother and child, especially during the time of actual observation. If the child made friendly or attention-getting overtures, the observer was to recognize him but to pay no further attention. As a rule, the child needed only this brief recognition to enable him to begin or return to an activity. If the child persisted, the observation was stopped and the child was told the observers were working and would talk to him when they had finished. The mother was usually helpful at such a time. In no instance did a child continue such overtures throughout an observation period.

The positions taken by the observers depended upon the setting, but they strove to maintain unobtrusive positions. If mother and child were in a relatively small room, the observers usually stationed themselves in or near the doorway, where they were apart from the immediate scene, could move out of the way if mother or child left the room, and could shift their position if the scene of interaction changed. When mother and child were in a larger room, the observers could again operate from the doorway or assume an unobtrusive position on either side or end of the room. When the interaction occurred outside the room, the observers remained in the background, as far as possible.

Whenever mother and child were out of each other's sight, the observers remained with the mother. This procedure was found to increase opportunities for observing mother and child interaction, since the child would return his attention to the mother, but the child, when the observer stayed with him, often interrupted the observer and made fewer contacts with the mother. This finding suggests again the value of a series of home observations in reducing the influence of the observer on the kind and quality of mother-child behavior.

RECORDING PRINCIPLES AND PROCEDURES

The behavior of child and adult is recorded individually for each 5-second period. Recordings are made in terms of category code letters (Table 1) and entered on a prepared scoring sheet, designed for 16 minutes of continuous recording, each square representing 5 seconds of time. A stop watch is used.

TABLE I

A SCHEDULE FOR OBSERVING AND SCORING CHILD-ADULT INTERACTION

	Scoring Symbol	
Category	*Adult*	*Child**
Nonattention	o	o
Neutral behavioral expression	o_0	o_0
Positive behavioral expression	o_1	o_1
Negative behavioral expression	o_2	o_2
Attentive observation	c	c
Recognition	e	e
Statement of condition or action		
Statement of situation	b_1	b_1
Statement of emotional state	b_2	b_2
Statement of action	b_3	b_3
Joint participation in activity	p	p
Offering information		
Verbal information	t_1	t_1
Information by demonstration	t_2	t_2
Information by demonstration and explanation	t_3	t_3
Giving help		
By assisting in completing task	h_1	h_1
By completing task	h_2	h_2
Reassurance	a	a
Seeking information		
Personal information	t_1—	t_1—
Impersonal information	t_2—	t_2—
Seeking help		
By requesting assistance	h_1—	h_1—
By requesting completion of task	h_2—	h_2—
Seeking reassurance	a—	a—
Seeking recognition	e—	e—
Seeking praise	m—	m—
Seeking affection	l—	l—
Seeking reward	g—	g—
Seeking permission	Q—	Q—
Orienting		
Boundaries	s_1	s_1
Time	s_2	s_2
Roles		
By status, power, age	$\mathrm{s}_{3\mathrm{m}}$	$s_{3\mathrm{m}}$
By indicating responsibility	$\mathrm{s}_{3\mathrm{n}}$	$s_{3\mathrm{n}}$

(*continued on next page*)

TABLE 1 (*continued*)

	Scoring Symbol	
Category	*Adult*	*Child**
Directing		
Suggestion	d_1	d_1
Command and threat	d_2	d_2
Command	d_3	d_3
Restricting		
By warning	r_1	r_1
By explanation	r_2	r_2
By threat of bodily punishment	r_3	r_3
By threat of loss of affection	r_4	r_4
By direct statement	r_5	r_5
By recognition and clarification	r_6	r_6
By limiting the use of objects	r_7	r_7
By physical restraint	r_8	r_8
Forbidding		
By explanation	f_1	f_1
By offering alternative	f_2	f_2
By recognition and clarification	f_3	f_3
By threat of bodily punishment	f_4	f_4
By threat of loss of affection	f_5	f_5
By direct statement	f_6	f_6
By removing object	f_7	f_7
By physical barriers	f_8	f_8
By physical restraint	f_9	f_9
By warning	f_{10}	f_{10}
Criticism		
Of person	n_1	n_1
Of person's production	n_2	n_2
Of person and person's productions	n_3	n_3
Disciplinary action		
By isolation	i_1	i_1
By deprivation of objects or space	i_2	i_2
By attacking person's objects	i_3	i_3
By corporal punishment	i_4	i_4
By deprivation of love or affection	i_5	i_5
Physical attack	v	v
Threat of attack		
By verbal expression		x_1
By gesture		x_2
By verbal and gestural means		x_3
Rejection		
By changing subject	w_1	w_1
By denying validity of statement	w_2	w_2
As a person	w_3	w_3
By ignoring or evading	w_4	w_4

(*continued on next page*)

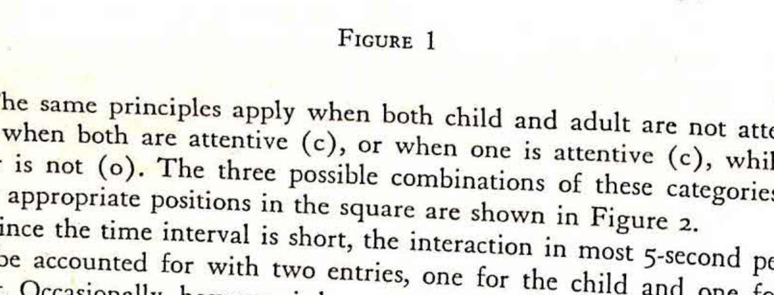
TABLE 1 (*continued*)

Category	Scoring Symbol *Adult*	 *Child**
Permission		
With qualifications	Q_1	Q_1
Without qualifications	Q_2	Q_2
Praise	m	*m*
Affection		
By physical means	l_1	l_1
By verbal means	l_2	l_2
By physical and verbal means	l_3	l_3
Reward	g	*g*
Cooperation		
Casual or straightforward	+1	*+1*
Enthusiastic	+2	*+2*
Compliance	/	*/*
Noncooperation		
With explanation	−1	*−1*
With alternative	−2	*−2*
Simple refusal	−3	*−3*
Strong refusal	−4	*−4*
Ambivalent statement	u	*u*
Interpretation		*z*
Restatement of content	z_1	
Verbalization of motor behavior	z_2	
Restatement of verbalized feelings	z_3	
Recognition of feeling in motor behavior	z_4	
Clarification of verbalized feelings	z_5	
Clarification of feelings in total behavior	z_6	
Association of current and past events	z_7	
Account of reality or translation of symbolic behavior	z_8	
Anxiety		
Little or none	0	*0*
Some	+	*+*
Much	++	*++*
Hostility		
Little or none	0	*0*
Some	−	*−*
Much	−−	*−−*

* Underlined on scoring sheet as necessary.

General Recording Procedures

As soon as the observation begins, the behavior of adult and child is noted, identified in terms of categories, and recorded by placing the symbols for the categories in appropriate positions in the interval square. These positions are rigidly defined and adhered to throughout the recording. Regardless of initiator and of stimulus or response qualities, *the first behavior to appear within a 5-second period is always entered in the upper left corner of the time period square.* The location in the square of behavior subsequent to this initial entry depends upon whether adult or child provides the initial behavior. If the adult provides the initial behavior, the child's behavior is recorded in the lower right corner of the square. (Note: Symbols referring to child's behavior were underlined on the scoring sheets as necessary.) For example: The adult suggests the child build with blocks (d_1), and the child responds enthusiastically (+2). This recording is shown in Figure 1 (*a*), which also illustrates another recording principle, namely, that *the lower right corner of the square is used solely for recording the child's behavior.* If the child provides the initial behavior within a 5-second period, the adult's behavior is always recorded above the square. For example, when the child directs the adult (d_3), and the adult cooperates in a straightforward way (+1), the interaction is recorded as in Figure 1 (*b*). That is, *only the adult's behavior can be recorded in the space above the square.* The same scoring principles apply when the initiating behavior of adult or child is obviously not responded to, as when the child ignores (w_4) the adult's question (t_2), which is recorded as shown in Figure 1 (*c*).



(*a*) (*b*) (*c*)

FIGURE 1

The same principles apply when both child and adult are not attentive (o), when both are attentive (c), or when one is attentive (c), while the other is not (o). The three possible combinations of these categories and their appropriate positions in the square are shown in Figure 2.

Since the time interval is short, the interaction in most 5-second periods can be accounted for with two entries, one for the child and one for the adult. Occasionally, however, it becomes necessary to make three recordings in one square. If the adult initiates the behavior in the record period, the third category entry will be the adult's and will appear above the square.

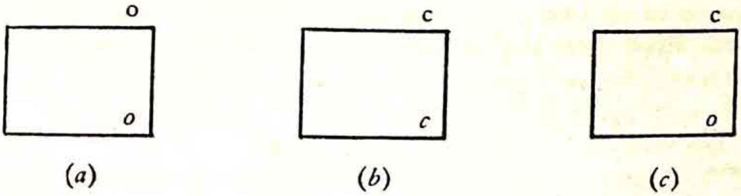


FIGURE 2

For example, the adult suggests that the child paint (d_1), the child asks about the brushes (t—), and the adult responds by telling the child where the brushes are (t_1). The recording of this behavior sequence is shown in Figure 3 (*a*). A similar procedure is followed when the child initiates the behavior. Figure 3 (*b*) shows the recording of a behavior sequence when the child asks a question (t—), the adult ignores it (w_4), and the child repeats the question (t—).

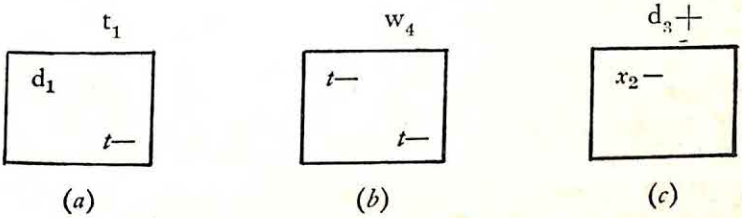
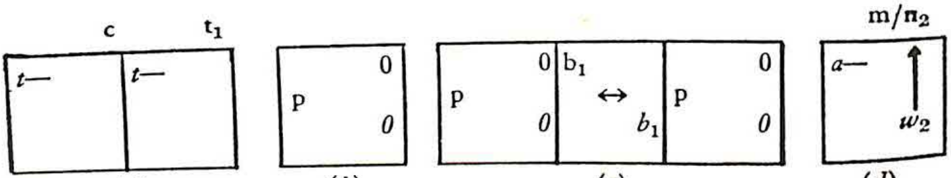


FIGURE 3

The previous illustrations have not included the anxiety-hostility rating, which actually accompanies each interactive category in the schedule. The symbol denoting the degree of anxiety or hostility immediately follows the behavior entry. For example, when the child, with an angry expression on his face, raises a hammer to strike the adult (x_2—), and the adult says, in an unsure voice, "Put the hammer down," (d_3+), the incident is recorded as shown in Figure 3 (*c*). Neutral, positive, or negative behavior expression in connection with nonattention (o or *o*) would be recorded as o_0, o_1, or o_2. For example, "child breaks car in solitary play" would be recorded as o_2.

When behavior extends beyond one 5-second period, it is recorded once every 5 seconds for as long as it continues. For example, when the child takes more than 5 seconds to ask a question (t—), and the mother listens during the first 5 seconds (c) and answers the question in the following

5-second period (t_1), a record is made as shown in Figure 4 (*a*). If, during observation of an interaction situation the child and adult separate and are out of visual range of each other, a dash is recorded in the time periods involved.



(*a*) (*b*) (*c*) (*d*)

FIGURE 4

Where it adds meaning to the recorded symbols, a short descriptive account of the behavior of adult and child is placed on the scoring sheet, directly above the square or squares concerned.

Special Recording Procedures

Category p, Joint Participation in Activity, differs from all other categories in representing simultaneous interaction; that is, it does not separate the ongoing interaction into specific individual behavior. Consequently, only one category entry is used to indicate the behavior of both adult and child. In this instance only, the anxiety-hostility rating is recorded for the adult in the upper right corner of the square, and for the child in the lower right corner, as shown in Figure 4 (*b*).

When a specific interaction extends beyond a 5-second period, especially if it is Joint Participation in Activity or Giving Help, another interactive behavior often occurs simultaneously. For example, while adult and child are engaged in working a puzzle (p), they may carry on a discussion completely unrelated to the puzzle (b). In this case it is necessary to record two kinds of behavior in one 5-second period, as shown in Figure 4 (*c*).

When child or adult makes a statement that contains two independent clauses, which represent behavior fitting two different categories, each aspect of the statement must be recorded separately. Suppose, for example, the adult uses constructive criticism. The child shows the adult a drawing of a boy, which he has just completed (*a*—). The adult says: "That's the best drawing you have made (m), but perhaps if it had some arms it would look more like a boy" (n_2). The child may respond to the praise or to the criticism. Assuming that he responds to the criticism, saying, "It does too look like a boy!" (w_2), the sequence would be recorded as shown in Figure 4 (*d*), with an arrow connecting the response and its initiator.

A final point may be made in relation to the need for information indicating whether behavior occurs in response to an expressed need or some other stimulus. Experience in using the schedule showed that such special subcategories were unnecessary, since the scores in the preceding time periods nearly always provide such information. For example, if Help (h) is given by the adult, and the symbol for Seeking Help, (*h*—), appears in the square immediately preceding, it would be evident that help was sought. If such a symbol does not appear, it may be assumed that help was given without being sought.

OBSERVER TRAINING

The observers' first task was to become thoroughly familiar with the details of the categorization system, their code letters, and the over-all scoring system. By extended discussion they arrived at common agreement concerning the specific types of behavior defined in the category system.

When the observers began to feel comfortable with the over-all details, they began observing the therapist and child from behind the screen. In the beginning they observed independently for one minute, then discussed each category entry that appeared within that time. An attempt was made to recall the specific kinds of behavior that gave rise to each recording, to analyze their disagreements, and to keep a record of the bases for them. Even when they agreed, the reasons for agreements were discussed, so that a clear understanding of the bases for scoring would result.

The bases for disagreements were then discussed among the authors of this paper. A further refinement of various categories was found necessary and was effected by more adequate definitions and division of categories, as a result of continuous interchange between authors and observers.

In view of the initial complexity of the task, the observers did not themselves keep time. A third person tapped out each passing time period and noted verbally a 30-second period, so that the observers could synchronize their observations. After the observers were familiar with the technique of recording, they began to practice timing themselves with a stop watch. Timing eventually became a rather simple task, for apparently a rhythm developed that made for accuracy. Nevertheless, periodic 30-second checks were made to assure synchronized recording.

After a month of intensive training, the observers were able to observe, more or less automatically, the complex parent-child or therapist-child interaction and enter the appropriate scores in the proper time blocks on the scoring sheet.

RELIABILITY

Reliability of observations was established by obtaining observer agreement on identifying behavior in appropriate time sequence. The following formula (4) was used:

$$\text{number of agreements} / (\text{number of agreements} + \text{number of disagreements})$$

In using a predetermined system of categories, in which the categories are recorded on a finely graduated time scale, several factors may operate to cause differences in the identification of behavior: the possibility of observers' seeing or hearing different events at the same time; imperfect familiarity with the categories; or an overly large number of categories in the schedule, making it impossible for any one observer to make reliable use of them.

It is sometimes impossible to determine whether disagreement between simultaneous records is a true difference in identification or only a timing error. When timing errors are possible, the computation of observer agreement becomes especially difficult. By inspection alone it is impossible to ascertain actual agreements and actual differences. To meet this situation we decided to establish certain arbitrary rules as to what constituted an agreement, a timing error, and an error attributable to other causes. When an identical behavior category or qualitative rating was recorded by each observer in the same time period, it was scored as agreement. A difference was scored as a timing error when the same category or rating was recorded by both observers in adjacent time periods. A difference that could not be attributed to timing error was considered a true disagreement.

Category Reliability

Category reliability refers to the ability of two observers to agree on identification of observed behavior. Individual category reliability measures for the three interactive situations appear in Tables 2 and 3. They are based on two 45-minute observation periods in each situation. Most of the categories show above 80 per cent agreement, indicating a reliability high enough to assure confidence in their use.

Only five categories had a reliability measure below 70 when computed on the total combined observations. Of these five, three were child categories (*Q*—, /, and w_1), and two were adult categories (z_3 and w_1). Each of these five categories had a total combined frequency of five or less and so required only one or two errors in agreement to yield low reliability. It is impossible to know whether these categories would prove reliable with increased frequency, but for practical purposes they can be accepted as functional.

Further observation is required to ascertain the specific reliability of categories that never appeared, the observers agreed, in the three situations observed. Such categories, quite possibly, would appear if observations were continued.

Observer Reliability

Observer reliability refers to the ability of two persons to agree upon the occurrence of an event in time. Contrary to the calculation of category reliability, timing errors are considered to be disagreements in calculating observer reliability. Only when an identical behavior category or qualitative

TABLE 2

RELIABILITY OF ADULT CATEGORIES

(Expressed in terms of frequency, and based on percentage of agreement for two 45-minute periods of observation in each of three adult-child interactive situations)

Category (Symbol)	T - C *Playroom* *fr.*	%	M - C *Home* *fr.*	%	M - C *Playroom* *fr.*	%	*All Obser- vations* %
o	3	25.00	397	89.92	30	100.00	91.51
c	1037	98.94	138	78.26	341	89.74	94.99
b	1	00.00	53	71.70	90	78.89	75.69
p	3	100.00	78	96.15	19	100.00	97.00
t_1	2	00.00	88	71.59	100	90.00	80.53
t_2	—	—	4	75.00	16	93.75	90.00
t_3	—	—	—	—	15	100.00	100.00
t_2—	3	00.00	104	75.96	73	95.89	82.78
h	19	78.95	28	96.43	4	100.00	90.20
s_2	5	100.00	—	—	—	—	100.00
s_{3n}	—	—	1	100.00	—	—	100.00
d_1	—	—	10	60.00	10	80.00	70.00
d_3	—	—	72	87.50	29	75.86	84.16
f_1	3	100.00	—	—	1	100.00	100.00
f_3	—	—	2	100.00	—	—	100.00
f_4	—	—	1	100.00	—	—	100.00
f_6	—	—	2	100.00	—	—	100.00
f_7	1	100.00	—	—	—	—	100.00
f_9	2	100.00	1	100.00	—	—	100.00
r_5	—	—	5	100.00	—	—	100.00
a	—	—	1	100.00	—	—	100.00
Q_2	—	—	1	100.00	1	100.00	100.00
Q—	—	—	1	100.00	—	—	100.00
m	—	—	7	85.71	4	100.00	90.91
n_1	—	—	2	100.00	—	—	100.00
l_2	—	—	1	100.00	—	—	100.00
l_3	—	—	10	90.00	2	100.00	91.67
z_1	15	60.00	11	72.73	7	100.00	72.73
z_2	6	100.00	—	—	1	100.00	100.00
z_3	7	28.57	—	—	1	100.00	37.50
z_4	11	90.90	—	—	1	100.00	91.67
z_5	5	100.00	—	—	—	—	100.00
+	—	—	3	100.00	2	50.00	80.00
−2	—	—	4	75.00	—	—	75.00
e	11	81.82	108	78.59	235	81.33	77.03
y	—	—	48	83.33	33	81.82	82.72
w_1	—	—	3	66.67	—	—	66.67

NOTE.—T = therapist; M = mother; C = child.

TABLE 3

RELIABILITY OF CHILD CATEGORIES

(Expressed in terms of frequency, and based on percentage of agreement for two 45-minute periods of observation in each of three adult-child interactive situations)

Category	T - C *Playroom*		M - C *Home*		M - C *Playroom*		*All Observations*
(Symbol)	*fr.*	%	*fr.*	%	*fr.*	%	%
o	1009	98.91	297	81.43	149	89.26	94.27
c	16	68.75	102	70.59	101	76.24	73.06
b	27	81.48	148	72.30	384	90.63	85.33
p	3	100.00	78	96.15	19	100.00	97.00
t_1	—	—	69	69.57	61	86.89	77.69
t_3	—	—	—	—	4	100.00	100.00
t_2—	—	—	58	77.59	64	87.50	82.79
d_3	9	77.78	20	85.00	15	100.00	88.64
s	2	100.00	1	100.00	4	100.00	100.00
h—	—	—	1	100.00	—	—	100.00
a	—	—	2	100.00	—	—	100.00
e—	—	—	31	83.87	—	—	83.87
Q_2	—	—	1	100.00	—	—	100.00
Q—	—	—	4	50.00	—	—	50.00
m—	—	—	—	—	3	100.00	100.00
l_3	—	—	1	100.00	—	—	100.00
l—	—	—	2	100.00	1	100.00	100.00
v_2	4	100.00	—	—	—	—	100.00
+1	1	100.00	25	80.00	31	80.65	80.70
+2	—	—	3	100.00	—	—	100.00
/	—	—	2	50.00	—	—	50.00
−2	—	—	2	100.00	—	—	100.00
e	25	76.00	112	79.46	71	76.06	77.88
w_4	29	75.86	56	82.14	37	78.38	79.51
w_1	—	—	—	00.00	—	—	00.00
w_3	—	—	—	100.00	2	100.00	100.00

NOTE.—T = therapist; M = mother; C = child.

rating is recorded by each observer in the same time period are the records considered to be in agreement.

Observer reliability was calculated for the categories, for the anxiety-hostility ratings, and for all entries within a time period, as shown in

Table 4. Observer reliability did not vary significantly in observations of mother and child in the home and in the playroom, or when these observations were compared with those of therapist and child. Reliability for the combined behavioral categories was 92.15; for the anxiety-hostility ratings, 99.10; and for the ratings within each time period, 88.51.

TABLE 4

OBSERVER RELIABILITY

(Expressed in terms of percentage of agreement on behavioral categories, anxiety-hostility ratings, and time period squares for two 45-minute periods of observation in each of three adult-child interactive situations)

	T - C *Playroom*	M - C *Home*	M - C *Playroom*	*All Observations*
Behavioral categories	96.41	87.58	92.72	92.15
Anxiety-hostility ratings	98.11	99.67	99.59	99.10
Time period squares	92.44	83.27	89.41	88.51

NOTE.—T = therapist; M = mother; C = child.

GENERAL TRENDS

This pilot investigation points to some tentative conclusions regarding adult-child interaction in the three types of situations studied.

In the frequencies of adult and child behavior (Tables 2 and 3), the behavior of adult and child tended to vary from one situation to another. However, there was greater similarity in the behavior of mother and child in the playroom and in the home than in either of these situations as compared with that of therapist and child in the playroom.

In the therapy situation, the range of behavior of both adult and child was comparatively limited. Approximately 93 per cent of the therapist's behavior consisted of attentively observing the child; that is, his most frequent response was in the nature of *being there,* watching and listening. Interpretation and Orientation were the only other approaches used to any considerable extent by the therapist.

The two children observed in play therapy had disturbing problems of severe anxiety and withdrawal. Their extreme inhibition in play necessarily limited the extent of interaction between child and therapist. These children were engrossed in solitary behavior approximately 95 per cent of the time. Nonrecognition or rejection of the therapist by ignoring or evading the therapist's stimulations was the most frequent interactive behavior, accounting for about 37 per cent of the child's total behavior. In frequency of occurrence, Statements of condition or action, Recognition of therapist's stimulations, Attentive observation, and Directing the therapist by command, followed, each accounting for 2 per cent or less of the child's behavior.

In comparison, the home and playroom behavior of mother and child tended to be interactive in nature rather than solitary. In both home and playroom the mother's behavior was characterized by Attentive observation, Statements of condition or action, Offering verbal information, Seeking impersonal information, and Recognition of the child's stimulation. Directing the child by command was used often by the mother in the home, relatively infrequently in the playroom. Mother and child participated jointly in activities considerably more often in the home (78 episodes) than in the playroom (19 instances). While the mother showed 397 instances of Nonattention in the home, she failed to follow the child's behavior in the playroom in only 30 instances. The mother occasionally forbade, restricted, and criticized the child in the home, but only rarely in the playroom.

The children's behavior with their mothers in both home and playroom (Table 3) shows frequent instances of Statements of condition or action, Nonattention, Recognition of mother's stimulations, Attentive observation, Offering verbal information, and Seeking impersonal information. Rejection of the mother by ignoring or evading her stimulations occurred more frequently in the home (56 times) than in the playroom (37 times).

These trends are presented as tentative and suggestive patterns rather than conclusive ones, and require examination and checking against larger samples and more comprehensive data.

Discussion

The schedule presents general purpose categories covering a wide range of adult-child social interaction. With such a research tool, comparable studies of adult-child interaction can be made, offering, for example, an opportunity to compare parent-child and teacher-child interactions.[11]

The schedule intentionally allows an objective description of ongoing behavior, without assumptions concerning the motivational determinants of such behavior. However, a study of the types of behavior descriptions obtained from the schedule indicates that the behavior can be interpreted according to any theoretical position.

The schedule includes what the authors assume to be the more important categories of adult-child social interactional behavior. The question of how many categories an observer could learn and use effectively was the guide in determining how many categories to include. We believe the number used in the schedule approaches the maximum for effective use; it includes, at any rate, the number that two observers could handle adequately.

---

[11] Two studies, one published, one in progress, involve use of the schedule. The first, concerned with therapist-child interaction, has been published (10). The second, comparing mother-child interaction in playroom and home, is the doctoral dissertation of Henry D. Schalock, working under the direction of Dr. Sigel at The Merrill-Palmer School and a graduate committee at the University of Nebraska, Department of Education, headed by Dr. Warren R. Baller.

Another problem concerns the reliability of all the categories. As shown in Tables 2 and 3, some categories never appeared in the sample observations. Conceivably, the appearance of some categories may require a very large sample. At this point we cannot delete the categories as useless, in view of the small sample, but must await a more extended use of the schedule.

The schedule has various limitations. Thus, in any preconceived category structure certain types of behavior will be omitted. In this schedule, for example, we are aware that much nonverbal social behavior, such as voice tone, gestures, postural signals, facial expressions, are omitted in favor of concentrating primarily on verbal behavior. Meaningful behavior of this kind might be the subject of further schedule construction. Another inherent limitation is that the schedule leaves the observer with a contentless description of the behavior observed. Verbal material is not available for the investigation of such questions as the symbolic significance of the response, language development or usage, and the like. Some other well-known limitations of such an observational procedure are problems of time sampling, the shortcomings of preconceived observation schedules, and the like.

Limitations notwithstanding, we offer this technique as useful for the observational purposes outlined. Further research should refine the technique further, and so increase its usefulness and extend its applicability.

SUMMARY

A schedule is presented for observation of adult-child interaction, with a total of 89 adult and 82 child categories, as well as anxiety-hostility ratings. The schedule has been tested and found reliable in studying parent-child interaction in home and laboratory (playroom) and in studying therapist-child interaction. Certain general trends, similarities, and differences in mother-child interaction in home and playroom and in therapist-child interaction in the playroom are noted. The general applicability, strengths, and weaknesses of the procedure are discussed.

REFERENCES

1. ARRINGTON, RUTH E. Time sampling in studies of social behavior: A critical review of techniques and results with research suggestions. *Psychol. Bull.*, 1943, 40, 81-124.
2. BISHOP, BARBARA MERRILL. Mother-child interaction and the social behavior of children. *Psychol. Monogr.*, 1951, 65, No. 11.

3. Collier, R. M. A scale for rating the responses of the psychotherapist. *J. consult. Psychol.*, 1953, 17, 321-326.

4. Festinger, L. *Analysis of methods of calculation of per cent agreement.* Laboratory Course in Child Psychology, 1944-45. (Mimeographed)

5. Finke, Helene. Changes in the expression of emotionalized attitudes in six cases of play therapy. Unpublished master's thesis, Univer. of Chicago, 1947.

6. Landisberg, Selma, & Snyder, W. U. Nondirective play therapy. *J. clin. Psychol.*, 1946, 2, 203-214.

7. Lebo, D. The relationship of response categories in play therapy to chronological age. *J. child Psychiat.*, 1952, 2, 330-336.

8. Moustakas, C. E. *Children in play therapy.* New York: McGraw-Hill, 1953.

9. Moustakas, C. E. The frequency and intensity of negative attitudes expressed in play therapy: A comparison of well-adjusted and disturbed young children. *J. genet. Psychol.*, 1955, 86, 309-325.

10. Moustakas, C. E., & Schalock, H. D. An analysis of therapist-child interaction in play therapy. *Child Develpm.*, 1955, 26, 143-157.

11. Osborn, D. K. Dependency in young children in relation to two maternal variables: A methodological study. Unpublished master's thesis, State Univer. of Iowa, 1952.

12. Seeman, J., *et al.* The development of empirical criterion measures in the study of play therapy. (Unpublished manuscript)

IDENTIFICATION WITH PARENTS AS A DETERMINANT OF DOLL PLAY AGGRESSION[1]

HARRY LEVIN and ROBERT R. SEARS
Cornell University *Stanford University*

A major problem in the study of personality development is to discover behavior variables that not only are relevant to the more important motivational systems of the child but which also permit precise measurement and some degree of experimental control. The so-called projective techniques have seemed promising in these respects. They provide responses belonging to significant motivational categories such as sex, aggression, dependency and achievement, but the responses are highly attenuated in their intensity and may be elicited safely and conveniently under controlled conditions. Projective responses have no special virtues other than these; they are not, in themselves, "true" measures of any "deep motives," nor do they have any unique validity as measures of personality qualities. They are simply acts that have presumably discoverable antecedents and that bear regular relationships to other acts. Fantasy behavior, in other words, is no different from any other behavior in its logical status as a reference event to be accounted for by principles of learning and action.

Experimental doll play has been rather widely used for the study of fantasy aggression in preschool aged children, both because this method requires less verbal skill than the visual-perceptual projective techniques and because it seems to be peculiarly effective in eliciting aggressive activity. Doll play offers an opportunity for measurement of three aspects of fantasy

---

[1] This article reports a selected set of findings from a larger study of "Identification in Young Children" conducted by the Laboratory of Human Development at Harvard University. Primary responsibility for the study rested with Robert R. Sears, Harry Levin, Eleanor E. Maccoby, Pauline S. Sears, and John W. M. Whiting. The study was made possible by a grant from the National Institute of Mental Health (M-461). Parts of this paper were presented at the 1953 meetings of the American Psychological Association in Cleveland, Ohio.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

aggression: (*a*) the *amount,* as indicated either by frequency of aggressive behavior units or by proportion of time devoted to aggression, (*b*) the *agents and objects* of aggression, i.e., who attacks whom, in thematic play with doll characters, and (*c*) the *quality* of the aggression in terms of such categories as directness, intensity, verbal or physical expression, and pro- or contra-social quality. These various attributes of aggression can be measured by the use of appropriate recording categories, and the resulting quantitative "scores" can then be used as standard reference events for which correlates can be sought. While these fantasy events are of considerable interest in their own right, they have added importance that comes from their, as yet, largely undiscovered relationships to comparable aspects of overt social aggressive behavior, and their inhibiting or facilitating influence on non-aggressive activity.

Previous researches have discovered some of the factors that influence these doll play measures. We will limit our consideration here to those that have been found to be related to the *amount* of aggression, since this is the measure with which the present study is concerned.

There are several aspects of the methods used for doll play that are relevant (10). Among these are (*a*) the amount and kind of interaction between the child and the experimenter, with a high level of interaction initiated by the latter being positively related to frequency of aggressive response; (*b*) the realistic quality of the doll house materials, the more realistic being favorable to aggression; (*c*) the duration of the play session, with a 20-minute session producing a relatively greater amount of thematic aggression than a full hour session, which tends to elicit a good deal of tangential behavior; and (*d*) degree of organization of the doll house and furniture, the more organized setup providing higher average aggression scores.

The actual handling of the child during his play session is also important. Experimental doll play is customarily done under conditions of high permissiveness for aggression, the experimenter adopting an interested but neutral attitude toward the child's behavior. With this procedure there is usually an increase in frequency of aggressive acts from the beginning of a session to its end, and also from session to session. When this permissiveness is replaced by what is perhaps a socially more customary attitude, one that involves reproof or disapproval for thematic aggression, the frequency of this kind of act decreases at once (3). It appears that the permissive attitude is anxiety reducing: it eliminates the constraint that the usual social control by adults ordinarily imposes on the child's aggressive activity, whether reality-oriented or fantasy. In sum, these methodological factors provide variations in the immediate instigating conditions.

There are other sources of influence, however, that are not subject to the experimenter's control. These are the child's relatively stable response potentialities that have developed either from his past experiences or are characteristic of his constitution. Some children are more aggressive than others in the doll play situation. In general, boys have been found to be more aggressive than girls (7, 8), accident-prone boys more so than non-accident-prone ones (5), and boys with their fathers living at home more so than those with their fathers absent from home (8, 12).

While it is probable that constitutional factors have some influence on these aggressive response potentials, our present concern is with the effects of differences in child-rearing experiences. The studies just mentioned are of some value in this respect, for they refer to groups which have been or could be examined with respect to the kinds of child training used by the children's mothers. For example, as has been shown elsewhere (11), there are important differences in the treatment accorded boys' and girls' aggression in early life, and these differences are relevant to the sex differences in expression of fantasy aggression.

There are many other differences between boys and girls, however, and it is difficult to assess the exact influence of all of them on doll play aggression. A more direct approach can be made by holding constant such gross sources of influence as sex of child, and searching for the effects of specific child-rearing practices as these are revealed in interviews with the mothers of children whose fantasy aggression is measured. This was the procedure used in an earlier investigation (13), which provided a crude assessment of the effects of one child-rearing variable on amount of fantasy aggression, and suggested the hypotheses that have been tested in the present research.

PREVIOUS FINDINGS

In the earlier study (13), 40 mothers of preschool aged children were interviewed about their child-rearing methods. From these reports, the mothers were rated on a scale that described the dimension *severity of punishment for the expression of aggression.* Their children were tested with two 20-minute sessions of doll play. Severity of punishment proved to be positively related to the amount (frequency) of doll play aggression. For the girls, the correlations were +.41 and +.46 for the first and second sessions respectively; for the boys, +.46 and +.23.

It is evident, of course, that this one child-training dimension was not accounting for a very substantial proportion of the total variance of the fantasy aggression scores. Another finding from the same study drew attention to the possible influence of *identification with parents.* The children were also observed in preschool, and measures were obtained of the amount of overt social aggression exhibited there. For the boys, the correlation of this measure with severity of punishment was +.56. But for the girls, the relation was curvilinear, with both high and low punishment associated with low overt aggression. It was as if the girls had been more severely punished on the average, although the ratings did not show this, and the highest punished ones had become inhibited and passive in their social expression of aggression. The hypothesis was suggested that girls, being more highly identified with their mothers, suffered more inhibitory effects from a given degree of severity of punishment.

HYPOTHESES ABOUT IDENTIFICATION

Our attention having been brought to the identification process, we considered what possible effects it might have on expressions of fantasy aggression. Its relevance seemed to lie, possibly, in the character of doll play as a situation that permits a child to play the role of whichever family-member he wishes. There is a presumption that he identifies in some degree with all the persons who both reward him and place demands on him; this includes all the members of his family, but preeminently the parents. In doll play he can perform these identified-with roles, and do the kinds of things that he understands those roles require. Thus, the higher the identification a child has with a parent, the more his doll-play behavior should be like the behavior that the child believes is characteristic of that parent.

In the present context, we will assume that aggressiveness is a quality in a parent's behavior that a child can recognize, and that the child can discriminate whether the parent shows much or little aggression in relation to other (nonaggressive) forms of action. Thus, if we also assume, as suggested above, that in doll play a child adopts the role of his most-identified-with parent and acts out that parent's characteristics, we arrive at our first hypothesis:

(*a*) The more strongly a child is identified with a given parent, the more nearly will he approximate in doll play the level of aggression he perceives as characterizing that parent. This means that if a boy is strongly identified with his father, and the father is very aggressive, the boy will display a high level of aggression in doll play. Conversely, if the father is relatively nonaggressive, the boy will be also. If, however, the boy's identification is weak, the father's level of aggressiveness should have little or no influence.

But with which parent does a child identify? By the age of five years, most children may be assumed to have developed their strongest identification with the same-sexed parent. This development is not the same for the two sexes, because the conditions for establishing primary identification are such that the mother is the initial identificand for both boys and girls. Only as the boy begins to perceive himself as a male, "like" his father and "unlike" his mother, can he begin to shift his identificand from mother to father. Hence, identification is more straightforward in the girl, and she develops a stable sex-typing more rapidly in this respect than the boy.

This latter assertion is supported by two findings from previous studies of doll play. If we accept the assumption that the doll most frequently chosen as agent of an act is the fantasy representative of the chief identificand, we can examine frequency of choice by children of both sexes. P. S. Sears reported (9) that the five-year-old girls of the present study used the mother doll most frequently for neutral positive (nonaggressive) acts, whereas boys used the mother and father dolls almost equally in the first session and used the father more only in the second session.

Another type of comparison was made by Johnson (4). She compared the relative frequency of pro-social and contra-social aggressive acts in five- and eight-year-old children of both sexes. The pro-social acts were chiefly disciplinary ones that represented *adult quality* of behavior. Both sexes increased sharply in this type of aggression from five to eight years of age, but the girls were very much higher than the boys at both ages, and the five-year-old girls were almost as high as the eight-year-old boys.

We can conclude from these findings that the chief identificand for a child at five to six years of age is most commonly the parent of the same sex, though the relative strength of such an object choice for identificand is greater for girls than for boys. This leads to our second hypothesis:

(*b*) The father's level of aggression will be most influential in determining the boy's frequency of aggressive doll acts, and the mother's level will be most influential for the girl's.

But what determines the child's perception of his same-sexed parent's aggressiveness? There appear to be two types of experience that are important, and their interaction produces a curious and complex difference between the sexes with respect to what child-training variables influence the frequency of doll play aggression.

The first of these two factors is perception of sex role. Characteristically the male is more aggressive. Among adults, there are more suicides, homicides and street fights by men. War has always been a male preoccupation. Rough contact games are mainly for males. Even in nursery school the sex difference is evident, in both overt and fantasy expression. The boy may be expected to perceive his father as a representative of males in general, and hence as relatively aggressive.

The other factor is the child's own *personal experience of parental aggression*. Even though a boy recognizes the aggressiveness of the male role, his own father may be relatively mild. The opposite variation can occur for a girl. For most children, it seems likely, the clearest measure of parental aggressiveness is derived from the severity with which the parents punish the child for his aggressive actions toward them. If his aggressive acts are met with vigorous counter-aggression—i.e., parental actions that the parents themselves view as punishment—the child will perceive parents as more aggressive than if his acts are met with only mild punishment. The important aspect of this experience lies in the distinctiveness of the cues it provides the child. Presumably, when the chief punisher is the parent with whom the child identifies more strongly, the cues will be more distinctive. Hence we get the following hypotheses:

(*c*) The frequency of doll play aggression will be greater in boys than in girls (because of the male role perception). This hypothesis has been verified by a number of studies (1, 7, 8).

(*d*) The frequency will be greater when punishment for aggression is severe than when it is mild. This was found to be the case in one previous study, mentioned earlier (13).

(*e*) Given severe punishment for aggression, the frequency of doll play aggression will be greater when the parent of the same sex is the chief disciplinarian.

If we now convert these five hypotheses to statements of what child-training antecedents will contribute to the frequency of children's doll play aggression, we find that the factors are different for the two sexes. For boys, we would expect the following for the reasons indicated: (*a*) *degree of identification* with parents, because the male role is an aggressive one, (*b*) *severity of punishment,* because this adds information for the boy's perception of his identificand's level of aggressiveness, and (*c*) *the father is the chief disciplinarian,* because this increases the distinctiveness of cues and spells out unmistakably to the boy that males are aggressive. In all likelihood, too, this increases the severity of frustration the boy suffers, because the punishment comes from the identificand, and hence adds to the boy's motivation for aggressive counter-actions.

For girls, the factors are different. Since the female role is *not* an aggressive one, the girl's identification with her mother should not in itself be a significant antecedent except as it may be associated with a mother who is a very aggressive person.

The antecedents should be (*a*) *degree of identification,* but only when there is high punishment, that is, the girl is provided with an atypical female model, (*b*) *severity of punishment,* because this provides an aggressive role model, but only when (*c*) *the mother is the chief disciplinarian.*

METHOD

The testing of these hypotheses required measurement of four variables. One was the criterion, or consequent, variable of frequency of aggressive acts in the doll play of five-year-old children of both sexes. The second was the degree of identification of these children with their parents. The third was the severity of punishment for aggression that the parents used on the children, and the fourth, an indication of which parent was the chief punisher of the child.

In connection with an extended study of the origins of identification in a group of 379 five-year-old children, these four measures were obtained. The last three were secured through interviews with the children's mothers.

The children were all enrolled in public school kindergartens in two suburbs of a New England metropolitan area. There were 202 boys and 177 girls, of various ordinal positions, and the families were fairly evenly spread from upper lower socio-economic status to upper middle. The mean age of the subjects was 5-7, the range being 4-11 to 6-7, with a standard deviation of 4.4 months. The details of this sample have been described elsewhere by Maccoby and Gibbs (6).

Doll Play

Each child was given two 20-minute sessions of doll play, never on consecutive days, but always within a five-day period. The play was scored in terms of behavior units with but two categories, aggression and neutral positive acts. Aggression was defined as an action that had the intent to irritate, hurt, injure, punish, frustrate or destroy a doll or equipment. It ranged in intensity from innocuous mischief (the girl sneaks out of bed) to severe accidents, fights and catastrophes (the father spanks the boy, the father has half his head sliced off, a bomb falls on the house and kills everybody). Verbal aggression such as scolding, threats, and uncooperativeness were included, as were instances in which discomfort was attributed to a doll (the boy is sick, sad, lost, etc.). The doll agent of each act was also recorded. The measure used in the present study is the "per cent aggression" for each session separately. This value is obtained by dividing the frequency of aggressive behavior units in a session by the total number of units of play. Nine experimenters collected these data. They had been trained to an adequate degree of comparability in their scoring, so their data have been combined. A full account of this doll play procedure is presented elsewhere (8).

Mother Interviews

The children's mothers were interviewed in their own homes by women interviewers who had been trained to use a standardized set of questions that covered the major areas of child rearing. This interview, which has been described elsewhere (11), consisted of 72 main questions and appropriate probe questions to assure that the necessary information would be obtained on each topic. The interviews were recorded and then were transcribed to typescript. They averaged between two and two and a half hours in length, and approximately 50 pages in single-spaced typed form. A series of rating scales and category codes was used for quantifying the interview information. Each interview was rated on every scale by two raters working independently. The reliabilities of the ratings were calculated by correlations between these independent judgments. The final score used for each scale on an interview was the pooled rating of the two raters; when the two ratings were seriously discrepant, the judges reexamined the interview and attempted to reach a closer agreement. Reliability coefficients were calculated *before* any such discussions.

From the interviews, measures were obtained of the three variables that were required for the present analysis. These, together with their Spearman-Brown corrected reliability coefficients, were (*a*) degree of identification with parental role that was displayed by the child ($r = .77$), (*b*) the severity of punishment given the child for expressing aggression toward the parents ($r = .77$), and (*c*) which parent was the usual punisher ($r = .95$). Some comments are needed to clarify the bases on which ratings on these three scales were made.

Identification

The rating of the child's degree of identification was based on the mother's replies to the following questions:

1. We'd like to get some idea of how X acts when he's naughty . . . when he has deliberately done something he knows you don't want him to do, when your back is turned; how does he act?
2. Does he ever come and tell you about it without your having to ask him?
3. When you ask him about something he has done that he knows he's not supposed to do, does he usually admit it or deny it?
4. What do you do about it if he denies something you are pretty sure he has done?

We reasoned that the identification process is necessary to the formation of super-ego, or conscience, and that the amount of identification the child has with parental roles can be estimated from the stage of development he has reached in internalized control. These four interview questions were designed to bring out the mother's observations of her child's behavior that were relevant to this matter.

The rating scale was labelled *evidence of super-ego in child,* and the raters were instructed to take into account the following: "High super-ego is inferred from confession of deviation, acting guilty and feeling bad over deviations, admitting rather than denying a fault. Hiding and other evidences of fear should be taken as evidence of *low* super-ego. The making of reparations is evidence of *high* super-ego. If any of the indicators of high super-ego are explicitly stated to have occurred because of fear, however, they should be discounted. For example, if child admits rather than denies because he has been severely punished for denying, admitting is less direct evidence of super-ego."

The following descriptive statements were used to define the five points on the rating scale; the per cent of children rated at each point is also given.[2]

---

[2] The two raters disagreed by one point, on this five-point scale, on 109 children. For purposes of this tabulation, half of the disagreed-upon cases were allocated to each of the adjoining scale points. For example, there were 35 children that both raters agreed should be "1" and 72 that both agreed should be "2." But there were 28 cases that were rated "1" by one rater and "2" by the other. Half of these, or 14 have been allocated to "1" and half to "2." For all data analyses, those cases have been defined as "1.5."

1. No evidence of super-ego. Child denies, hides, doesn't seem unhappy when naughty.	13%
2. Slight super-ego.	28
3. Moderate super-ego. May not confess directly, but looks sheepish; seldom denies.	38
4. Considerable super-ego. May come and express affection to mother when he has been naughty; tries to repair damages.	17
5. High super-ego. Feels miserable when naughty; always confesses, never denies; strong need for forgiveness.	3
Not ascertained.	1
Total	100%

The following example will indicate the type of report that was rated toward the high end of the super-ego scale:

> I: We'd like to get some idea of how Ed acts when he's naughty; when he deliberately does something he knows you don't want him to do when your back is turned, how does he act?
>
> S: Very seldom does that. But a few times he has done something that he shouldn't do, that I don't know anything about, if I'm in the other room, he just can't hold it in very long and finally comes in to me and he says, "Mother," and I'll say, "What?" "I did something I shouldn't have done." Instead of leaving it and getting away with it he usually comes over and tells me what he's done. He usually comes, I mean, and it's not very long after he's done it. He can hardly hold it in to himself, you see.

One anecdote that was volunteered by a mother, although not necessarily an example of high super-ego, suggests the way in which a child can display adult attitudes and play the adult role; it shows how a highly identified child takes over some normally parental behavior:

> S: I've taught him never to touch a medicine cabinet. I showed my little boy bottles—medicine bottles—and I've told him—he was sick, oh, about a year ago, in the hospital; and he's never forgotten it, and I don't think he ever will. He's told his brothers about it plenty of times. And I told him, "If you took your medicine bottle, you're going to get sick like you did then." He won't go near that medicine cabinet. And if he sees one of his brothers going to touch it, he'll tell him, "You'll go to the hospital, and you'll stay in the hospital for a long, long time; and you won't see Mommie or Daddy. And you won't have no toys to play with—just lay in bed all the time and get needles." And now none of them will go up there. I think he's really the one that broke them of going to the medicine cabinet.

Severity of Punishment for Aggression Toward Parents

An extended report of this scale and of the influence of such punishment on overt aggressive behavior in the home is published elsewhere (11). The scale was a five-point one that took account of how severe the punishment was that the child actually received, regardless of who did the punishing.

Which Parent Was the Usual Punisher

The mother was asked, "Who usually disciplines, you or your husband?" The replies were categorized as "husband," "wife," or "about equal." Of the 379 cases, 108 were judged to fall in this last category, and 11 could not be evaluated. Since the hypotheses with which we are concerned here relate only to the condition in which the preponderance of punishment is by one parent or the other, we have limited the data analysis to the 260 children whose mothers reported either "husband" or "wife" as the usual punisher.

Since there were a few cases which could not be evaluated on the *identification* and *severity of punishment* scales, the total number of children actually utilized for the present analysis was 241, of whom 126 were boys and 115 were girls.

TEST OF THE HYPOTHESES

The various hypotheses outlined above offered certain predictions as to what combinations of the three mother interview variables would be associated with high or low frequency of doll play aggressive acts. These predictions have been tested by a triple classification analysis of variance, separately for the sexes. Except for *who usually punishes,* which contained two mutually exclusive categories, the variables were continua; it was necessary, therefore, to break each distribution into as nearly equal halves as possible for the analysis of variance. It was impossible to secure exactly equal numbers of cases in the eight required cells, however, and hence the correction for unequal cells as given by Walker and Lev (14) was used. The analysis was done separately for each of the two doll play sessions.

Boys

The analysis for boys is shown in Table 1, which gives the mean % Aggression scores for the eight groups of children representing the eight possible combinations of the three antecedent variables. The table also shows that in both sessions there were two major sources of variation in the % Aggression scores. One of these was *identification* and the other an interaction between *identification* and *agent of punishment.* The influence of these sources of variation is clearly shown in the means; the two highest means for the first session are "Father punishes, High Identification." On the second session, two of the three highest were in the same category.

The effects of these two variables can be more easily seen when the data are regrouped, as in Table 2. When the mother punished, the degree of identification had no influence on the amount of aggression displayed, but when the father punished, high identification produced high aggression.

TABLE 1

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS, AND ANALYSIS OF VARIANCE ACCORDING TO IDENTIFICATION, SEVERITY OF PUNISHMENT FOR AGGRESSION, AND AGENT OF PUNISHMENT (BOYS ONLY)

	HIGH IDENTIFICATION				LOW IDENTIFICATION			
	Severe Punishment		Mild Punishment		Severe Punishment		Mild Punishment	
	Father	*Mother*	*Father*	*Mother*	*Father*	*Mother*	*Father*	*Mother*
Session I	18.6	13.9	23.0	13.8	5.8	12.1	10.2	13.6
Session II	28.8	20.1	25.9	26.7	12.4	23.9	18.7	23.4
N	18	11	18	13	14	18	14	20

	Session I				Session II			
Source	*df*	*Sum of Squares*	*F*	*p*	*df*	*Sum of Squares*	*F*	*p*
Agent of Punishment	1	2.21	.19	n.s.	1	8.61	.52	n.s.
Severity of Punishment	1	13.00	1.10	n.s.	1	11.28	.68	n.s.
Identification	1	95.22	8.08	<.01	1	66.70	4.04	<.05
Agent × Severity of Punishment	1	6.85	.58	n.s.	1	.91	.05	n.s.
Agent × Identification	1	69.62	5.91	<.05	1	72.60	4.39	<.05
Severity × Identification	1	.32	.03	n.s.	1	.55	.03	n.s.
Agent × Severity × Identification	1	.32	.03	n.s.	1	33.21	2.01	n.s.
Residual	118	1389.84			118	1950.91		

These findings confirm the hypotheses that place emphasis on the boys' identification with an aggressive male role model as a determinant of the frequency of aggression in doll play.

One hypothesis is not supported, however. The severity of punishment for aggression proved to have no influence on doll play aggression. This finding is in contrast to that from the earlier investigation, although in that one the correlation of this variable with second session doll play

TABLE 2

MEAN % AGGRESSION IN TWO DOLL PLAY SESSIONS ACCORDING TO THE DEGREE OF IDENTIFICATION AND THE AGENT OF PUNISHMENT (BOYS ONLY)

	High Identification		Low Identification	
	Father Punishes	*Mother Punishes*	*Father Punishes*	*Mother Punishes*
Session I ...	20.8	13.8	8.0	12.9
Session II ..	27.3	23.7	15.6	23.6
N	36	24	28	38

aggression was quite small in boys (.23), while the girls' was larger (.46). Since neither *identification* nor *agent of punishment* was measured in that study, it is possible that the small positive correlation was a product of some unknown correlation with these variables.

We have no certain explanation for the lack of influence of severity of punishment, on boys, in the present study. One possible reason may be that, by the age of five years, most boys have been so much exposed to aggressive male models that the added experience of observing the father's being aggressive has little or no added effect. Possibly boys get enough training in male role aggressiveness that the only factors required for making them behave aggressively themselves are a strong identification with the adult role and the clear cues that derive from paternal punishment.

TABLE 3

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS, AND ANALYSIS OF VARIANCE ACCORDING TO IDENTIFICATION, SEVERITY OF PUNISHMENT FOR AGGRESSION, AND AGENT OF PUNISHMENT (GIRLS ONLY)

	HIGH IDENTIFICATION				LOW IDENTIFICATION			
	Severe Punishment		Mild Punishment		Severe Punishment		Mild Punishment	
	Father	*Mother*	*Father*	*Mother*	*Father*	*Mother*	*Father*	*Mother*
Session I	8.8	13.2	11.0	5.7	5.2	9.5	8.1	7.8
Session II	15.4	20.1	12.3	8.5	7.7	11.3	13.6	10.4
N	9	18	9	24	11	13	11	20

	Session I				Session II			
Source	*df*	*Sum of Squares*	*F*	*p*	*df*	*Sum of Squares*	*F*	*p*
Agent of Punishment	1	1.20	.16	n.s.	1	.21	.10	n.s.
Severity of Punishment	1	2.10	.28	n.s.	1	11.76	1.06	n.s.
Identification	1	8.20	1.20	n.s.	1	22.11	1.99	n.s.
Agent × Severity of Punishment	1	25.56	3.42	$<.07$	1	29.26	2.64	$<.10$
Agent × Identification	1	3.00	.40	n.s.	1	.03	.00	n.s.
Severity × Identification	1	5.28	.71	n.s.	1	48.51	4.37	$<.05$
Agent × Severity × Identification	1	3.25	.43	n.s.	1	.36	.03	n.s.
Residual	107	800.30			107	1187.90		

Girls

The case is quite different with girls. As Table 3 shows, the interaction between *severity of punishment* and the *mother-as-agent* of punishment approaches statistical significance in both doll play sessions. The data are rearranged in Table 4 to show the influence of these two variables. When the mother was the punisher and the punishment was severe, the fantasy aggression was greater than under any other combination of variables.

TABLE 4

MEAN % AGGRESSION IN TWO DOLL PLAY SESSIONS ACCORDING TO THE SEVERITY OF PUNISHMENT FOR AGGRESSION AND AGENT OF PUNISHMENT (GIRLS ONLY)

	Severe Punishment		Mild Punishment	
	Father Punishes	*Mother Punishes*	*Father Punishes*	*Mother Punishes*
Session I . . .	6.8	11.7	9.4	6.6
Session II . .	11.2	16.4	13.0	9.4
N	20	31	20	44

This finding supports the prediction that exposing girls to an aggressive role model increases their doll play aggression. It is to be noted that the converse holds true as well; when the mother was the punisher, but punishment was mild, doll play aggression was least. In other words, the mothers who did not punish severely permitted their daughters to take over the socially modal nonaggressive female pattern. Severely punitive mothers erected potential barriers against their daughters' copying of at least this one aspect of approved femininity.

The role of identification appears to have been of some importance in this connection. In the second doll play session, there was a significant interaction between identification and severity of punishment. The data are regrouped in Table 5 to show the joint effects of these two predictor variables. Girls who were both highly identified and severely punished exhibited more aggression than the girls falling in the other three combinations. These latter did not differ from one another, on the average, since the remaining influential variable (agent of punishment) varied randomly among the groups.

This finding adds further support to the prediction that identification will lead to much aggression for girls only when it is associated with an atypically aggressive female role model, as exemplified in a severely punishing mother.

TABLE 5

MEAN % AGGRESSION IN THE SECOND DOLL PLAY SESSION ACCORDING TO THE DEGREE OF IDENTIFICATION AND THE SEVERITY OF PUNISHMENT FOR AGGRESSION (GIRLS ONLY)

	High Identification		Low Identification	
	Severe Punishment	*Mild Punishment*	*Severe Punishment*	*Mild Punishment*
Session II . .	18.5	9.5	9.6	11.5
N	27	33	24	31

Comparison of the Sexes

There are two striking sex differences in the antecedents of doll play aggression frequency. One is with respect to identification and the other to severity of punishment. In boys, identification is a significant positive influence, both alone and in interaction with other variables. In girls, it is apparently less so. There are two possible explanatory factors to be considered. One is that the measure of identification is of less discriminating value with girls than with boys, because girls are developmentally farther along in the process. As was mentioned earlier, boys presumably develop identification with their mothers first and then must shift to fathers for the appropriate role model. Girls, in other words, may have reached an asymptotic level of identification in their sixth year, while the boys of this age are strung out along a fairly wide range of identification. Thus this variable would assume greater significance for boys.

The other factor has to do with severity of punishment. The male role is an aggressive one, and the actual day-to-day demonstration of this fact by the father's use of severe punishment may add little or nothing to the boy's understanding of how he should behave. It is more important for the boy that he have high identification with the role. For girls, however, the appropriate sex role model is normally nonaggressive. Only when the particular model most visible to the girl (her mother) is especially aggressive does the girl adopt an aggressive way of behaving.

There is one further hypothesis we might suggest for further research on this problem of sex differences. Among the children in this sample, the boys were subjected to more *physical punishment,* according to the mothers' reports (11). If the aggressive behavior in doll play were scored as to whether it was physical or verbal in content, we would therefore expect that the boys would show more physical and the girls more verbal expressions. This sex difference was found in studies by P. S. Sears (8) and by Bach (1), but no child training antecedent measures were available. In the present study, the differentiation of content was not scored.

Identification with the Aggressor

Extended clinical study of children led Anna Freud to the discovery of a defense mechanism she has termed identification with the aggressor (2). This mechanism is a technique by which a child can reduce his fear of an aggressor, perhaps a punitive parent, by identifying with that person and adopting his aggressive behavior.

The findings of the present study are fairly well in accord with what one would predict from this mechanism if one assumes that frequency of doll play aggression is a measure of the aggressive motivation derived from that identification process. It is the girls who have most to fear from their mothers, i.e., the daughters of severe punishers, who show the highest aggression. The same does not hold for the boys, but for both sexes it is the children who are ordinarily punished by the *parent of the same sex* who show greatest aggression. The connection between the measure of identification and amount of aggression may equally well be interpreted either as evidence that identification with the aggressor had taken place or, as we have done above, as an indication of the requirement of identification in order that agent of punishment be effective. In other words, the rating of high identification in those children (especially boys) who showed high aggression could be viewed as an *effect* as well as a *cause*.

We are inclined to rely on the interpretation we presented in the introduction to this paper rather than on the conception of identification with aggressor because we do not have any way of determining, from the present data, whether the fantasy aggressions had anxiety-reducing effect. This is an essential assumption underlying an interpretation in terms of this defense mechanism. As the data stand, however, there is no factual reason why the doll play aggressive acts cannot be interpreted as evidence of the operation of this mechanism.

TABLE 6

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS, WITH CORRELATIONS AND DIFFERENCES BETWEEN SESSIONS AND SEXES

	N	Session I	Session II	r	t
Boys	202	14.28	21.84	.61**	3.30**
Girls	177	8.64	12.76	.69**	2.25**
t between sexes		4.32**	4.95**		

** Significant at $p < .01$.

NORMATIVE FINDINGS

Since standardized doll play as used in this study is a common method of research with children, it is useful to catalog norms of such behavior. An analysis of the influence of sex, age, ordinal position and father's absence on doll play aggression has been reported by P. S. Sears (8). The present results, where applicable, will be compared with hers.

Sex Differences

As has been stated earlier, the boys devoted significantly more of their actions to aggression in both doll play sessions than did the girls (Table 6). This confirms other findings (1, 4, 7, 8) and may be taken as a reasonably stable characteristic of doll play behavior.

Session Changes

For both sexes, there was a significant session-to-session increase in the proportion of aggression units (Table 6). This finding is also in line with previous studies (3, 8) and has been interpreted to mean that the deliberate permissiveness of the experimental procedure progressively acts to reduce the inhibitions about aggression that the child has heretofore acquired.

Although there was a session-to-session increase in the mean % Aggression scores, there was a correlation between sessions of .61 for boys and .69 for girls. P. S. Sears (8, p. 35) reported session-to-session correlations of .74 for five-year-old boys and .53 for five-year-old girls, with 25 cases in each category.

Ordinal Position

P. S. Sears (8, p. 24) found that, for both sexes, children who were "youngest" in their family or who were "only" children showed more aggression than "oldest" children. In the present study, four ordinal positions were analyzed: (*a*) *only* children, (*b*) *oldest* of two or more children, (*c*) *youngest* of two or more children, and (*d*) *middle* of three or more children. The results are presented in Table 7.

TABLE 7

MEAN % AGGRESSION FOR TWO DOLL PLAY SESSIONS ACCORDING TO THE SEX AND ORDINAL POSITION OF SUBJECTS

	BOYS			GIRLS		
	N	Session I	Session II	*N*	Session I	Session II
Only	33	19.36	30.64	19	8.88	13.76
Oldest	59	13.08	19.32	50	8.76	12.84
Youngest	62	14.68	18.72	51	8.48	13.36
Middle	48	13.28	22.84	57	7.84	11.76

The present findings only partially replicate the earlier ones. In both doll play sessions, "only" boys exhibited the most fantasy aggression, but among the girls, there are no sizeable mean differences in % Aggression which are attributable to ordinal position.

Descriptive categories such as ordinal position or socio-economic status are useful in understanding behavior insofar as they lead the researcher to investigate the particular patterns of experiences which people occupying such categories have had. Other analyses of the child-training data upon which the present study is based indicate that "only" children were judged more strongly identified with their parents, as measured by the super-ego rating, than were children of other sibling statuses (11). This finding, together with the effect of identification on aggression for boys as reported above, makes reasonable the result that the "only" boys were most aggressive. One can speculate that P. S. Sears may have had some highly identified boys in her "younger" group, but the evidence is not available.

Socio-economic Status (SES)

Previous studies of doll play aggression have utilized children from widely differing social class groups. Those done in university (1, 3, 7, 13) or private preschools (4) have drawn subjects from mainly middle-class populations. P. S. Sears reported that her group (8, 12) represented "a predominantly upper-lower and lower-middle class position." Yarrow (15) had a definitely lower-class group. In the absence of any direct comparison of the aggressiveness of children from the different SES levels, it has been difficult to know how widely the results of the various studies could be generalized.

The present sample was chosen to represent a wide range of SES. Warner's Index of Status Characteristics was computed for each family on the basis of the father's income and occupational status. Maccoby and Gibbs (6) have reported in detail the rationale and computation of the score, and the SES characteristics of the present sample. For our purposes the sample was divided into two groups which may be described as "working class" and "middle class." As may be seen from Table 8, there were no differences in mean % Aggression scores attributable to SES in either doll play session.

TABLE 8

MEAN % AGGRESSION: THE INFLUENCE OF SOCIO-ECONOMIC STATUS

	BOYS			GIRLS		
	N	Session I	Session II	*N*	Session I	Session II
Middle Class	117	13.76	22.36	86	8.80	12.58
Working Class ..	82	15.28	22.48	88	8.76	13.88

SUMMARY

A test was made of the hypothesis that the frequency of five-year-old children's fantasy aggression is in part a function of the children's identification with aggressive role models. The frequency of aggression was measured in two sessions of doll play with each of 241 children, 126 boys and 115 girls. The degree of the children's identification with their parents, the severity with which the parents punished aggression, and an indication of which parent usually did the punishing were ascertained from interviews with the children's mothers.

As was predicted, boys who were highly identified, and who had the distinctive cues for male aggressiveness provided by being usually punished by their fathers, showed the highest frequency of aggression. Identification was related to high aggression in the girls only when it was associated with severe punishment by mothers who usually did the punishing, i.e., when the girl was identified with an aggressive role model.

The present study replicated several earlier reports in finding that, in both doll play sessions, boys were more aggressive than girls, and that, for both sexes, there was an increase in the amount of aggression from the first to the second session. "Only" boys were more aggressive than boys of other ordinal positions, but there were no ordinal position differences for girls. The difference for boys was related to differences in child training practices by the parents of the "only" children. No differences in aggression were attributable to the socio-economic status of the children's families.

REFERENCES

1. BACH, G. R. Young children's play fantasies. *Psychol. Monogr.*, 1945, 59, No. 2.
2. FREUD, ANNA. *The ego and mechanisms of defense.* London: Hogarth, 1937.
3. HOLLENBERG, ELEANOR, & SPERRY, MARGARET. Some antecedents of aggression and effects of frustration in doll play. *Personality,* 1951, 1, 32-43.
4. JOHNSON, ELIZABETH Z. Attitudes of children toward authority as projected in their doll play at two age levels. Unpublished doctor's dissertation, Harvard Univer., 1951.
5. KRALL. VITA. Personality characteristics of accident repeating children. *J. abnorm. soc. Psychol.,* 1953, 48, 99-107.
6. MACCOBY, ELEANOR, & GIBBS, PATRICIA. Methods of child-rearing in two social classes. In W. E. Martin & Celia B. Stendler (Eds.) *Readings in child development.* New York: Harcourt, Brace, 1954.
7. PINTLER, MARGARET H., PHILLIPS, RUTH, & SEARS, R. R. Sex differences in the projective doll play of preschool children. *J. Psychol.,* 1946, 21, 73-80.
8. SEARS, PAULINE S. Doll play aggression in normal young children: influence of sex, age, sibling status, father's absence. *Psychol. Monogr.,* 1951, 65, No. 6.
9. SEARS, PAULINE S. Child-rearing factors related to playing of sex-typed roles. *Amer. Psychologist,* 1953, 8, 431 (Abstract).
10. SEARS, R. R. Influence of methodological factors on doll play performance. *Child Develpm.,* 1947, 18, 190-197.
11. SEARS, R. R., MACCOBY, ELEANOR, & LEVIN, H. *Patterns of child rearing,* in press.

12. Sears, R. R., Pintler, Margaret H., & Sears, Pauline S. Effect of father separation on preschool children's doll play aggression. *Child Develpm.*, 1946, 17, 219-243.

13. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, Pauline S. Some child rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-234.

14. Walker, Helen, & Lev, J. *Statistical inference.* New York: Henry Holt, 1953.

15. Yarrow, L. J. The effect of antecendent frustration on projective play. *Psychol. Monogr.*, 1948, 62, No. 6.

THE DEVELOPMENT OF PREJUDICE: A CRITIQUE OF THE *HOROWITZ FACES TEST*

EUGENE E. LEVITT

Institute for Juvenile Research, Chicago[1]

Twenty years ago Horowitz (10) published the results of a study of the development of anti-Negro prejudice in children using a "faces" test, which he devised for the purpose. The Horowitz Faces Test consists of a cardboard about 12 inches square on which are the faces of 12 boys, eight Negroes and four whites. The ages probably range from about 5 years to about 12 years. Each face is identified by a letter from A to L. The subjects are required to perform two tasks: (1) rank all the faces for preference; (2) respond to 12 sociometric questions, such as "Show me all those you want to be in your class at school," "Show me all those you like," and "Show me all those you want to sit next to in the movies."

The score on the first task, the Ranks, is the sum of the rankings of the four white boys. The possible range is from 10 (the four white boys preferred) to 42 (the four white boys least preferred.) On the second task, the Show Me, the score is the percentage of white boys of all those selected.

The original Horowitz study has never been replicated, and in fact, the present writer knows of only three researches in which the test has been used. Nonetheless, Horowitz based certain important conclusions on his work and a survey of the literature since 1935 indicates that these conclusions have been widely accepted by social scientists.[2] Briefly, the conclusions

---

[1] This paper was written while the author was at the Iowa Child Welfare Research Station, University of Iowa. Its preparation was partly supported by Research Grant MH-301 from the National Institute of Mental Health, U. S. Public Health Service. The author is indebted to Drs. Maurice S. Schaeffer and Boyd R. McCandless for many excellent comments and criticisms which have influenced the exposition.

[2] Horowitz' conclusions, especially the first one, are quoted in many texts in social psychology, for example, (3, 13, 16, 22, 29). They constitute the basis of a section on the development of prejudice in a recent volume on the subject (21). The first conclusion has been quoted by Hyman and Sheatsley (12) in support of criticisms of the California studies of the anti-democratic personality.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

were: (1) Attitude toward the Negro is determined not by contact with the Negro himself, but by contact with the prevalent attitude toward the Negro. (2) Prejudice develops gradually, the growth curve being of the sigmoid type. (3) With increase in age from 5 to 14, an individual's attitude becomes more integrated, i.e., more consistently expressed despite differences in measuring instruments.

The purpose of the present paper is to examine critically the original Horowitz study, especially in light of other researches and findings. The critique is subdivided into four main sections concerned with (1) experimental design and statistical analysis, (2) test validity, (3) test reliability and (4) developmental theory.

DESIGN AND ANALYSIS

The 725 subjects used by Horowitz ranged from kindergarten to grade eight, and included small samples of elementary school children from urban Tennessee and Georgia and from rural Georgia. Four hundred and seventy were from New York City schools. Sixteen had Communist parents, and 35 attended a racially-mixed school. All were boys. The experimental design is, of course, greatly simplified by having a unisexual population. It precludes the necessity for two faces tests, one for each sex, and reduces the data collection and analysis by half. Unfortunately, it also limits the conclusions, unless one accepts, *a priori,* that there will be no sex differences worth noting. In view of the studies (cf. 24, 25) which show differences in racial and religious tolerance between adult males and females, it would hardly seem prudent to make such an assumption.

The four appendices to Horowitz' article (10) contain means, correlation coefficients, and critical ratios, but they are devoid of variance data. However, the size of the means relative to the ranges appear to indicate that the distributions of both the Ranks and the Show Me scores are markedly skewed. For the 472 New York City children, the mean Ranks score is 14.9 and the range is 10–42. The Show Me has a mean of 65 and a range of 0–100. Estimated standard deviations of the two distributions for the New York sample can be derived from Horowitz' Table 1 which is concerned with regression equations and curve-fitting. The sigmas are 20.7 for the Ranks and 20.6 for the Show Me.

It is most likely that a distribution of 470 scores with a mean of 14.9, a sigma of 20.7, and a range from 10 to 42 is not normal. In all probability, the Show Me distribution is also not normal. All of the statistical techniques used by Horowitz are parametric, depending for their proper application upon underlying normal distributions. Hence they have been improperly applied. They include the various correlation coefficients, critical ratios, and comparison of curves, in short, the fundament of Horowitz' primary conclusions.

The use of inapplicable statistical techniques is a serious defect, but it should not be construed as reflecting unfavorably on the experimenter in this case. Except for chi square and rank correlation, distribution-free statistical methods were not available in 1935. Horowitz had little choice but to go ahead with the use of known techniques.

Nonetheless, we must regard Horowitz' conclusions with caution. For instance, the conclusion that attitude toward the Negro is not a function of actual contact with the Negro himself is based primarily on statistical comparisons which show that the Southern samples do not differ from the New York City sample. The actual critical ratios are not given except for the Ranks for the sixth grade. It is possible that the variability of the Southern samples is considerably less than that of the Northern sample. We do not know, but if so, a Type II error resulting from the use of a parametric comparison becomes a definite possibility. A nonparametric procedure like the median test might show a significant difference between the two distributions.

The possibility of a Type II error is heightened by the small number of cases in the Southern samples. The largest subsample is the fourth grade urban Georgia group which numbers 35. The smallest is the kindergarten Tennessee group with only 4 subjects. Of the 12 subsamples, only 3 contained more than 20 subjects. The total Tennessee sample comprises 60 subjects spread over six grades. Even with proper statistical procedures, the number of Southern subjects is too small for any definite conclusions.

VALIDITY

If a subject were to rank the 12 faces in a random fashion, his Rank score would be 26, the midpoint of the range. If he selected a given number of faces on the Show Me in the same way, his score would be 33·33, since the white faces are one-third of the total number on the card.

A significant deviation from these chance scores indicates that some biasing factor or factors are operating. If the deviation is in the direction of a lower score than 26, or a higher percentage than 33·33, the biasing factor is race prejudice, according to Horowitz. A chance score indicates that the subject is "responding without prejudice," while a deviation above 26 on the Ranks and below 33·33 on the Show Me means that the subject is "responding with a preference for Negroes."

In attitude scaling, a high score is usually taken to mean that the subject is prejudiced. A low score means that he is unprejudiced, while a score somewhere in the middle of the distribution indicates an average amount of prejudice. There is no score which corresponds to the "preference for Negroes" although any of the low scores could be so interpreted. Most researchers in the field of race relations are not concerned with preference for Negroes as distinct from lack of prejudice. There are several good reasons for this unconcern. One reason is that while there are many unprejudiced persons, there are probably very few white individuals who generally *prefer* Negroes to whites. This applies to children as well as adults. The present writer knows of no actual data on this point, but it appears to be a highly plausible assumption.

A peculiarity of the Faces Test is masked by the expression "a preference for Negroes." Since a low Ranks score and high Show Me score indicate prejudice against Negroes, then a high Ranks - low Show Me score actually shows a prejudice *against whites.* If one desires to hold fast to the preferences for Negroes idea, then its opposite is logically preference for whites, rather than prejudice against Negroes.

An expression like prejudice against whites (and preference for Negroes) is obviously almost meaningless when applied to a group of whites. As pointed out previously, the social scientist is disinterested in such a concept, and has no reason to try to measure it. Yet, its existence and measurement are inextricably bound up with the assumptions basic to the Faces Test.

If we simply accept the assumption that prejudice against whites is a highly improbable phenomenon in a group of white children, we can explain Horowitz' results without concluding that his group means indicate anti-Negro prejudice. Assuming that anti-white prejudice is nonexistent in the group, a Ranks score higher than 26 or a Show Me score lower than 33.33 (within certain limits of error, of course) should not occur except by chance. These two scores constitute the no-prejudice limit of the test. The two other limits are 10 for the Ranks and 100 for the Show Me. The effective ranges are, therefore, 26–10 for the Ranks, and 33.33–100 for the Show Me. The mid-points are approximately 18 and 67, and we would expect the empirical means to fall somewhere around those two scores. As it turned out, the respective means for 674 unselected Northern and Southern subjects in Horowitz' sample were about 15 and 66, which correspond fairly closely to the hypothetical averages. When viewed in this light, the Faces Test shows that little or no prejudice exists in the sample tested by Horowitz.

Following along with this analysis, what would be the proper conclusion if the empirical mean actually was 26 or 33.33? This would mean that large numbers of subjects are manifesting anti-white prejudice. Since this is *a priori* improbable, the test must be at fault. In the absence of any extenuating evidence, we would have to conclude that the test is invalid. Horowitz' assumptions are completely twisted; the results which he calls indicative of prejudice are actually those to be expected purely by chance, while a realization of what he feels is the chance expectation would nullify his test.

Suppose we retrace several paragraphs and assume for the moment that Horowitz' assumptions are tenable. A group which yielded the mean scores to be expected by chance is unprejudiced. It must also be expected that the scores of this group will be distributed in some way around the respective means. The theoretical chance distributions are easily determined. For the Ranks, for instance, there are 495 possible scores[3] falling symmetrically about the mean of 26. The sigma of this distribution is 5.89. If we assume that the group has responded to the test in a purely random fashion, this is approximately the distribution to be expected. Statistically speaking, this distribution includes the scores of individuals who are probably prejudiced, viz., the 12 scores falling more than two sigmas below the mean (less than 14.22). That is, the fact that the *group* is unprejudiced does not in this case preclude the possibility of *prejudiced individuals* within the group. Since the group is ranking the faces randomly, we cannot be sure just which individuals are the prejudiced ones. Assuming adequate test reliability, a retest should yield the same mean score, but individual scores might change considerably. However, any subject whose personal chance distribution includes a score two sigmas below the mean ought to be suspect. At any rate, it is reasonably sound statistically to say that an individual who scores more than two sigmas below the mean is prejudiced. Indeed, it is a necessary statement if any individual prediction is ever going to be attempted.

Now suppose that we obtain a distribution with a mean of 18 and a sigma of 1.1. The mean differs significantly from 26, but the distribution does not include a single score lower than 14.22. Conclusion: the group is prejudiced, but it does not have a single prejudiced individual in it! It is obvious that we must insist that the variances of the empirical and the theoretical distributions be roughly the same if this paradox is to be avoided. But this is not likely if the mean is 18 or thereabouts. The lower limit of the range is 10 which would leave base-line room for less than 1.5 sigma below a mean of 18. If the sigma is 5.89, the distribution is going to quite skewed. A parametric comparison of the obtained and theoretical means is hence inappropriate. A satisfactory nonparametric comparison like the median test might not reveal a significant difference since it is probable that the medians of the obtained and theoretical distributions are closer together than the means.

The use of deviation from chance as a criterion of test validity is thus questionable. Even if this method stood up under logical analysis, the validity of the Faces Test as a measure of prejudice would be far from demonstrated. The test results might differ from chance due to the functioning of a biasing factor, but it would still remain to be shown that this factor was what we call race prejudice. The Faces Test directly measures *preference* for whites over Negroes. This alone is not commonly regarded as bigotry. Logically it is quite a leap from preference for, to prejudice against. Preference for one kind of stimulus does not necessarily imply hostility toward a second kind. As Allport (1) points out, the concept of hostility is a basic one in a definition of race prejudice. Horowitz made no real attempt to demonstrate that responses to the Faces Test were motivated by hostility toward the Negro. He recorded spontaneous comments of three children, one Northern and two Southern, to show that they were "recognizing the racial aspect of the tests," (the Northern child is a 4-year old girl) and 7 unclassified comments indicating that the children were "responding realistically." Even if there were a sufficient number of such remarks by representative members of the experimental population, their merit as evidence for validity would be doubtful.

---

[3] The theoretical distribution of Ranks scores is obtained by means of a formula for all possible combinations of 4 ranks out of 12 ranks. This formula is 12! / (4!) (8!).

In 1950, Mussen (18) published the results of a study using the Faces Test as a measure of anti-Negro feeling. He reported (17) that there was a correlation of .52 between Faces scores and choice of cabin-mates by white boys in a racially-mixed summer camp in the Northeast. Mussen derived a single Faces score by arbitrarily weighting the Show Me score five times and combining it with the Ranks score by means of standard scores, a procedure which practically eliminates the Ranks from consideration. The relationship between the Faces Test and choice of cabin-mates indicates that the former is valid as a sociometric indicator for Northern boys. The general results (18) are in accord with theoretical considerations, and we may accept Mussen's findings as indicating a degree of validity for the Faces Test. This is, however, the only such instance in the literature.

The Mussen study is one of the three published correlational researches in which the Faces Test was utilized as an index of prejudice. The others (14, 15) report negative relationships between the Faces Test and other measures. Levitt and McCandless (14) reported that both the Ranks and the Show Me were essentially unrelated to two conventional attitude scales for the measurement of anti-Negro feeling in children. McCandless and Holloway (15), using the Mussen combined score, found no correlation between the Faces Test and six measures of intolerance of ambiguity. The samples in both studies consisted of boys and girls in intermediate elementary school grades in two medium-sized Midwestern towns. It is possible, of course, that the Faces Test might be related to attitude or intolerance of ambiguity measures in populations in other sections of the country. In any event, the absence of relationships (which would be predicted to exist on theoretical grounds) in at least one kind of population limits the generality of conclusions stemming from the Faces Test.

Both Mussen (17) and Horowitz (11) are of the opinion that the Faces Test should produce different results than an ordinary attitude scale because of differences in the test structure, and because each test is likely to be tapping a different subarea of a nonunitary trait. This explanation could account for a low positive correlation between the Faces Test and the attitude scale, but it falls short of explaining an absence of relationship. Different subareas of a characteristic ought to be somewhat related, otherwise there is no empirical basis for regarding them as part of the same characteristic.

Certain other findings indicate that the preference for whites expressed on the Faces Test need not spring from hostility toward Negroes. It is reasonable to expect that children, with their burden of insecurity in a complex world, will prefer the familiar so long as it is unthreatening. None of the faces of Horowitz' test are personally known to the subject, but white skin is more usual in his private milieu, and hence more familiar in the general sense. That characteristic which most resembles himself becomes an important basis for preference where other bases are lacking.

Several studies support this reasoning. Springer (27) has shown that preschool children of three different racial groups manifested a modal preference for faces of their own race. There were no differences in the degree to which each group preferred its own faces. In a study by Clark and Clark (4) it was found that identification with a Negro or a white face by Negro preschool children was a function of the lightness or darkness of the responding child. Fifty-six per cent of the very dark children, but only 36 per cent of the light-skinned Negro children, identified with the Negro face. This seems to be a clear-cut instance of preference for the self-resembling face, regardless of other considerations.

The Clark and Clark experiment also gives another indication that preference on a faces-type test is probably unrelated to hostility. As the Clarks note, the responses of the light-skinned children suggest that the identifications were made "not on the basis of socially defined racial group differences, but on the basis of physical characteristics. . . ." (4, p. 160). The data from the Springer study (27) are in accord with the Clarks' conclusion. When asked for reasons for preference, Springer's children gave physical characteristics in 74 per cent of the instances. "Looks like me" made up more than 9 per cent of the reasons, while nationality was referred to in only 3 per cent of the reasons. Helgerson (9) found that when subjects and pictures of both sexes were used, the sex of the pictured face is more important than race in determining choice of prospective playmates.

It is true that all these studies used preschool children as subjects while Horowitz' subjects were grade school pupils. Furthermore, it is possible that an overtly expressed reason may be at variance with underlying motivation. Nonetheless, these findings are highly suggestive. They indicate that it is quite possible that children's preferences for faces are based on similarity to themselves, or upon more superficial characteristics than race *per se*. If this is true, it seems unlikely that the preference for white faces manifested by white subjects responding to the Horowitz Faces Test is motivated by hostility toward the Negro faces.

Some additional light on this point is cast by an analysis of the valences of the individual faces in Horowitz' test. By designed controlling of albedo, the Faces Test is arranged so that there are four light-skinned Negro faces and four dark-skinned Negro faces. In the alternate form of the test, the albedos are reversed, i.e., the light-skinned faces are dark, and the dark-skinned faces are light. By comparing group preferences for individual faces on the two forms, Horowitz concluded that preference for the lighter Negro faces exists only at the kindergarten and first grade levels.

However, an analysis of Faces Test data from a sample of 76 boys tested by Levitt and McCandless (14) fails to confirm Horowitz' finding. The approach was an analysis of variance into individual degrees of freedom using the method described by Snedecor (26). Using this procedure it is possible to determine whether or not there are significant differences between the light-skinned and the dark-skinned Negro faces each taken as a group. This analysis was done for both the Ranks and the Show Me.[4] For the Show Me, the data were the absolute number of times the face was selected per subject.

The resulting F-ratios are 7.25 for the Ranks, and 9.60 for the Show Me, both significant beyond the .01 level. In each case, the light-skinned group is the preferred one. The mean Ranks scores are 6.69 for the light-skinned group and 7.42 for the dark. The respective Show Me scores, i.e., mean number of times chosen per subject, are 5.57 and 4.81. These results appear to indicate rather definitely that in this sample of fifth grade subjects, albedo of the Negro faces is a factor determining preference.

An interesting side light is revealed by the analysis into individual degrees of freedom. In addition to the differences already noted, the light-skinned Negro faces, the dark-skinned Negro faces, and the white faces all differ significantly *among themselves* for both the Ranks and the Show Me. That is, some of the faces in each group are preferred to others in the same group. If we rank the mean scores per face on each test, a rank order correlation of .94 is obtained, indicating that the relative valence of each face is consistent from test to test.

The unequivalence of faces within groups raises the question of possible bias due to the stimuli themselves. An expressed preference of some kind for Negro face B (mean Ranks score, 4.58, mean Show Me score, 7.86) may mean something quite different from a preference for Negro face G (mean Ranks score, 8.54, mean Show Me score, 4.10). The same consideration applies to white face L (mean Ranks score, 4.72, mean Show Me score, 7.96) and white face F (mean Ranks score, 6.57, mean Show Me score, 6.19). If we accept the basic hypothesis that prejudice can be measured by preference for faces, then we would expect that an individual who selected the single Negro face, B, would be more prejudiced than one who chose the single Negro face, G. Rejecting face L would indicate less prejudice than rejecting face F. The usual method of scoring the Faces Test does not consider such differences, but simply treats all of the white faces, and all the Negro faces as equivalent. It might be added parenthetically that if the Faces Test is internally consistent, we would expect a high degree of consistency of response to individual faces. This may be true of the Show Me, but a rough inspection of the data indicates that it will hardly hold for the Ranks. For instance, face L, which has lowest (most preferred) mean Ranks score of the whites, was ranked above all the other white faces by only 18 of the 76 subjects.

---

[4] Technically speaking, Snedecor's procedure is not appropriate for ranked data. The writer knows of no method for breaking down the variance of ranked data into individual degrees of freedom. However, we may assume that Snedecor's method gives an approximate result.

The Show Me score is simply the percentage of white faces of those selected, irrespective of the absolute number of white faces chosen. The highest possible Show Me score, 100, could be obtained by choosing a single white face and no Negro faces, or by choosing 48 white faces (the maximum possible) and no Negro faces. It seems indisputable that a subject following the latter procedure is the more "prejudiced" of the two, though the test will not distinguish between them. We ought to find, then, that the tendency to select white faces is inversely proportional to the tendency to select Negro faces. The more white faces a subject selects, the more prejudiced he is, hence he should select fewer Negro faces.

This adds up to a negative correlation of some degree between the number of white and the number of Negro faces chosen. Actually, Levitt and McCandless (14) found just the opposite. For the sample of 76 Midwestern fifth grade boys, there was a *positive* correlation of .49 between the number of white faces and the number of Negro faces chosen on the Show Me. For a comparable sample of 78 girls, the correlation was .58.

This seems to indicate that responses to the Show Me are determined to some extent by a factor which we might call "accepting strange faces." This factor may, or may not, confound the Show Me results, but at any rate, it is something different from "accepting Negro faces," which was evidently assumed to be the key factor underlying responses to the Show Me.

A common method of establishing the validity of a new test involves the use of contrasting criterion groups. Suppose it is known or hypothesized on good theoretical grounds that the members of Group A have more of variable X than the members of group B. If the test shows, in fact, that this is the case, we conclude that it has some validity as a measure of variable X. Contrariwise, if Group A does no better, or worse, than Group B, we are led to suspect that the test is invalid as a measure of variable X. Of course, the test constructor can always fall back on a stubborn operationism, and insist that not only is his test valid, but that his results demonstrate that the supposed superiority of Group A is an "oft-repeated cliché" with no foundation in fact. The tenability of his claim will rest largely on the soundness of the evidence supporting Group A's superiority.

Not many social psychologists or sociologists would deny that there is a greater prejudice against the Negro in the South than in the North. As Horowitz himself once put it: "Most attempts to analyze attitude toward Negroes as a single, generalized unit suggest that Southern samples are least friendly, most 'prejudiced,' most reactionary, least tolerant, that the Northern samples are the reverse" (11, pp. 206-207). It is quite possible that Horowitz originally used the contrasting Northern and Southern groups in his study in an effort to show that the Faces Test could discriminate between them, viz., that the Southern group would show more prejudice. When it turned out that the groups did not differ (we shall assume for the moment that proper statistical procedures would lead to the same conclusions as improper ones) Horowitz faced alternate conclusions: (1) that the Faces Test is invalid as a measure of anti-Negro feeling; (2) that Southern white children are really not more prejudiced against Negroes than Northern white children. Obviously, he chose the latter, and surprisingly, many social scientists went along with him.

Horowitz used other contrasting groups in an effort to demonstrate test validity, and to support his conclusions. His small sample of children of Communists have a considerably higher Ranks mean and a lower Show Me mean than the other Northern white boys. The purpose of the inclusion of this group is obvious: Communists have a more favorable attitude toward minority groups than other people. Since that is what the test shows, in fact, then the test is valid. But the contrasting groups kind of validity, as pointed out previously, depends upon the soundness of evidence supporting the hypothetical contrast. It will be agreed, no doubt, that Communists *manifest* a greater acceptance of Negroes, but it is far from certain that this acceptance stems from genuine tolerance. As Shils (23) has pointed out in his penetrating analysis of the California studies of the anti-democratic personality, the leftist, like the rightist, is characterized by a rigid adherence to a particular set of stereotyped beliefs. The underlying personality structures may be similar, though the clichés differ. The Communist who conforms to his code *must ostensibly* accept minority group members, even though he seethes with suppressed hostility toward them. This prescribed credo of tolerance makes the Communist sample a poor one for contrasting groups validity.

Thirty-five sixth grade subjects from a mixed school were included in Horowitz' testing. Eleven of these were white children. Their mean Faces Test scores do not appear to differ from those of the children in the remainder of the New York City sample.

Horowitz therefore regarded this finding as further evidence that anti-Negro feeling is not determined by contact. The contrast is enervated by the small N and inappropriate statistical analysis, but more importantly, it conflicts with expectation based on a number of other researches. Negro children formed about 70 per cent of Horowitz' class. Criswell (5), using a sociometric technique, demonstrated that a few Negro children in a predominantly white class may be regarded as pets, while a large Negro representation is seen as a distinct threat. We would expect, therefore, that Horowitz would have found more anti-Negro feeling in this class than in others if his test was a valid index. One can argue that it is Criswell against Horwitz, but other data support the former's findings. For example, in the small-town Midwestern population sampled by Radke and Sutherland (19), prejudice did not appear until about age 11 or 12. The Negro population in the community was tiny. In the samples from large, Northeastern urban areas (where the proportion of Negroes is much higher) tested by Trager and Yarrow (28) and by Goodman (6), there was an appreciable smattering of anti-Negro feeling at the preschool level. Forty per cent of a New England sample (2) reported a first experience of anti-Negro feeling by age 11, while only 20 per cent made the same report in South Dakota (20).

This is another instance of the use of contrasting groups, and there are the same possibilities. Such evidence as there is indicates that the white children in the mixed class ought to show more hostility. According to the Faces Test, they do not. Either the assumption is incorrect, or the test is invalid. It appears that the scales tip in favor of the latter conclusion.

RELIABILITY

Mussen (17) reports an estimated split-half reliability of .94 for the Show Me. No coefficient is given for the Ranks, for some unknown reason. Horowitz does not report any statistical estimates, but gives the correlations by grade between two forms of both tests, and test-retest correlations for four grades after a six-months' period. The average correlation between the two forms is .68 for the Ranks, and .79 for the Show Me. The respective average test-retest coefficients are .59 and .50.

Since little data on the second form of the tests are offered, it is not possible to evaluate the correlation between forms. The test-retest correlations are, of course, extremely low, but this may be a function of the unusually large interval between testings. Especially with children, such an hiatus is apt to attenuate the coefficient.

The reliability found by Mussen appears satisfactorily high. However, the high split-half correlation may result in part from the fact that a sizeable proportion of subjects have selected all or nearly all 144 faces in response to the 12 Show Me items. In the Levitt-McCandless experiments, almost 10 per cent of the boys chose all 144 faces, while an additional 10 per cent chose 131 faces or more. Naturally, the split-half reliability of this 20 per cent of the results is going to be very high, probably .90 or better, and will markedly influence the reliability in the whole group. It is also possible that the percentage of Mussen's sample choosing all 144 faces was greater than in the Levitt-McCandless work.

As a possible check on the figures given by Horowitz and Mussen, split-half reliabilities were computed from the scores of the boys in the Levitt-McCandless study. The Ranks coefficient was obtained by correlating the sum of the rankings for the white faces C and F, with the sum of the rankings for white faces I and L. The Guttman coefficient (7) for these data is .48. The reliability for the Show Me is .89, an estimate obtained by correlating the percentage of white faces chosen on the first six questions with that on the last six. The reliabilities obtained from the Levitt-McCandless data are approximately the same as the Show Me coefficient found by Mussen and the test-retest reliability for the Ranks reported by Horowitz. We may conclude, at least tentatively, that the reliability for the Ranks is around .50, for the Show Me, about .90.

There is a connection between the Ranks and Question 12 of the Show Me—"Show me all those you like"—which permits an interesting and unusual investigatiton of reliability. Since the Ranks test is a ranking of the faces for liking, it ought to "scale," so to speak, with Question 12. That is, the faces which are selected in response to Question 12, should be found consecutively ranked from the top on down in the Ranks, with no gaps. A subject who chooses faces B, D, F, G, and K on Question 12 should also place these five faces in the first five rankings, the order being immaterial. It would be inconsistent, to say the least, if a subject ranked any face which he did not like ahead of those which he did like. In this sense, the Ranks and Question 12 are equivalent stimuli.

Analysis of test results show that inconsistency occurs more often than not. Of the 76 fifth grade boys, 42, or 56 per cent, had rankings inconsistent with Question 12. Of the 34 whose rankings were consistent, 25 had selected all of the faces in response to Question 12, (that is, it was impossible for their rankings to be inconsistent). Of the remaining nine, three had selected 11 of the 12 faces, which does not leave much room for inconsistency.

In other words, the stated preferences show a marked unreliability, even though the interval between presentations of the equivalent forms is only a few minutes. This raises a question of the significance of the choice of a face by a grade school child. Can the child actually be expected to express a meaningful preference among individuals merely by observing a face? It might be argued that inconsistency *per se* does not make for unreliability of the test as a measure of anti-Negro feeling. This argument is likely to be specious since in all but eight of the 42 inconsistent rankings, the deviation in rankings from the selections of Question 12 involved changes in position of a white face or faces relative to the Negro faces. That is, in each of these 34 instances, the Ranks score would change if it were made consistent with Question 12.

The lack of equivalence of the Ranks and Question 12 together with the relatively low Ranks reliability suggest that the Faces Test is not a particularly reliable instrument.

DEVELOPMENTAL THEORY

Horowitz developed three "theoretical growth curves of attitude" from his data.[5] These curves are the basis of his conclusion that prejudice develops gradually, following an S-shaped pattern. They are reproduced in Figure 1.[6] In selecting the curve for the Show Me as the representative

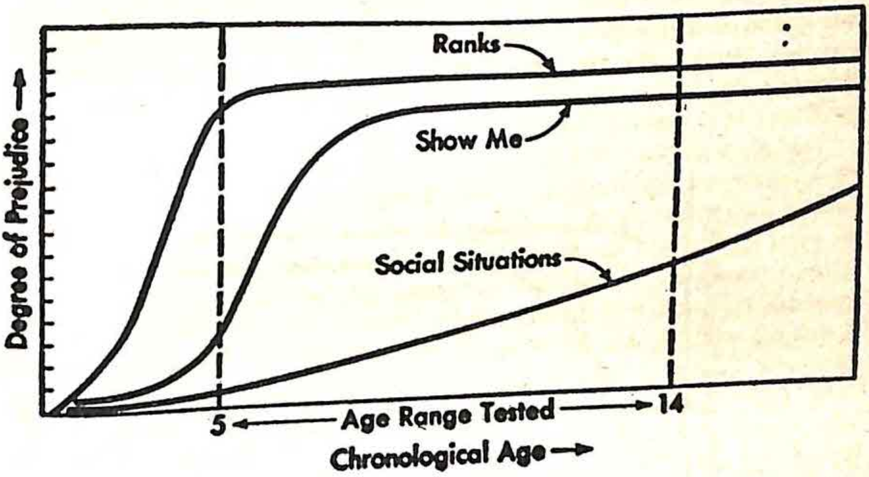


FIGURE 1—Theoretical growth curves of attitude toward the Negro according to Horowitz.

pattern, Horowitz says that "the three curves derived in this study may merely represent segments from different parts of three S-shaped curves" (10, p. 25). This speculation is, of course, a possibility, but the curves in Figure 1 are somewhat misleading. The actual data begin at the dotted line marking off age 5. The parts of the curves to the left of the dotted line are extrapolations, not based on obtained data. Figure 2 shows the curves stripped of extrapolations. It can be seen that there is nothing sigmoid about any of the curves, and that, of the three, only the Social Situations has a gradually rising slope. (Even the little slope at the beginning of the Ranks curve is unwarranted according to Horowitz' analysis of the data.)

---

[5] The third curve represents the Social Situations Test, a pictorial-type instrument similar in design and intent to the Faces Test. It was also used in Horowitz' original study (10). It is more or less independent of the Faces Test having an average correlation of .16 with the Ranks and .23 with the Show Me. Its test-retest reliability over the six-months interval was only .16.

[6] Figure 1 is a reproduction of a slightly modified version of Horowitz' original figure. It is taken from the recent text by Hartley and Hartley (8, p. 704).

The conclusions concerning development which appear warranted by the curves are: (1) Attitude as measured by the Ranks shows no real development. The attitude is completely developed at age 5, and does not change in succeeding years. (2) Attitude as measured by the Show Me develops rapidly from age 5 on, and is completely developed by age 8 or 9. After that, it does not change. (3) Attitude as measured by the Social Situations develops gradually from age 5 on, is still developing at age 14, and may conceivably develop further after that age. There is no statistical justification for choosing any one of the three as the dominant pattern, nor does there seem to be any logical way in which the three curves can be fused.

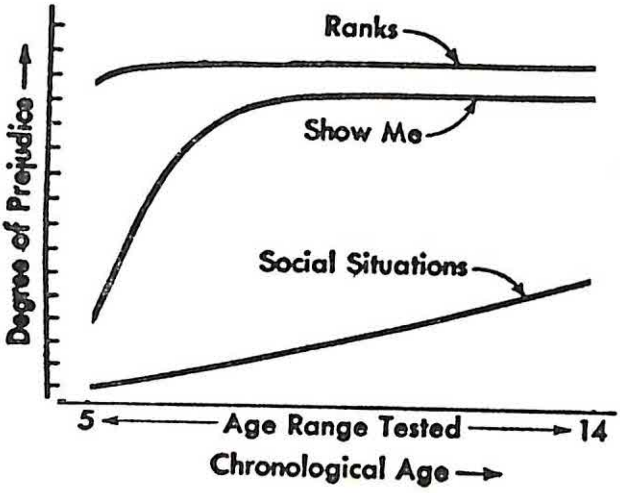


FIGURE 2—Horowitz' theoretical growth curves with extrapolations removed.

While it is true there is no statistical reason for looking upon one of the three curves as the curve, the Social Situations test has the most in its favor rationally. It is the only one which shows a continuing development over the grade school period. The Show Me curve could conceivably begin to ascend again after age 14, but it does not seem logical that there should be an attitudinal vacuum for the five preceding years.

In view of the defects of the test and the faulty statistical procedures, one cannot make any statement concerning the development of prejudice based on the Horowitz Faces Test. To be on fairly safe grounds, a conclusion concerning development should be based on a reasonably valid and reliable test, a cross-sectional sample of both sexes from various parts of the country, and proper statistical methodology. Such bases are not yet available, so far as the present writer knows.

Summary

The Horowitz Faces Test and the original work with it (10), which have been the primary basis for significant conclusions concerning the development of anti-Negro prejudice in children, were critically examined. Defects in research design and analysis, principally the use of inappropriate parametric statistics, were found. Analysis indicates that validity of the Faces Test is unsatisfactory for a number of reasons: (1) The use of deviation from chance as evidence of prejudice is logically untenable. (2) The test measures preference for whites or Negroes, but it has not been shown that these preferences spring from hostility. Contrary evidence is presented by the writer. (3) Two of three correlational studies report no relationships between the Faces Test and tests to which a valid prejudice measure should relate on theoretical grounds. (4) The various faces of the test are not equivalent stimuli, though the scoring of the test assumes that they are. (5) An unexpected positive correlation between number of Negro faces and number of white faces chosen has been found. (6) Evidence from contrasting groups used by Horowitz to support his conclusions is questionable.

Reliability of the Ranks was found to be unsatisfactory, though reliability of the Show Me was adequate. It was shown that subjects tended to respond differentially to the Ranks and to a very similar stimulus in the Show Me. The question of the meaning of a preference for a face by a child is raised.

Further analysis indicates that Horowitz' conclusions concerning growth of attitude are unsupported by his own data. None of the obtained curves is actually sigmoid as he claimed.

It is concluded that in view of the various shortcomings of the Faces Test and Horowitz' experiment, no definite statements concerning the development of prejudice are warranted.

References

1. Allport, G. W. *The nature of prejudice.* Cambridge, Mass.: Addison-Wesley, 1954.
2. Allport, G. W., & Kramer, B. M. Some roots of prejudice. *J. Psychol.,* 1946, 22, 9-39.
3. Bird, C. *Social psychology.* New York: Appleton-Century, 1940.
4. Clark, K. B., & Clark, Mamie K. Skin color as a factor in racial identification of Negro preschool children. *J. soc. Psychol.,* 1940, 11, 159-169.

5. CRISWELL, JOAN H. A sociometric study of race cleavage in the classroom. *Arch. Psychol.,* 1939, No. 235.
6. GOODMAN, MARY E. *Race awareness in young children.* Cambridge, Mass.: Addison-Wesley, 1952.
7. GUTTMAN, L. A basis for analyzing test-retest reliability. *Psychometrika,* 1945, 10, 255-282.
8. HARTLEY, E. L., & HARTLEY, RUTH E. *Fundamentals of social psychology.* New York: Knopf, 1952.
9. HELGERSON, EVELYN. The relative significance of race, sex, and facial expression in the choice of playmates by the preschool child. *J. Negro Educ.,* 1943, 12, 617-622.
10. HOROWITZ, E. L. The development of attitude toward the Negro. *Arch. Psychol.,* 1936, No. 194.
11. HOROWITZ, E. L. Race attitudes. In O. Klineberg (Ed.), *Characteristics of the American Negro.* New York: Harper, 1944.
12. HYMAN, H. H., & SHEATSLEY, P. B. The authoritarian personality—a methodological critique. In R. Christie & Marie Jahoda (Eds.), *Studies in the scope and method of the authoritarian personality.* Glencoe, Ill.: Free Press, 1954. Pp. 50-122.
13. KATZ, D., & SCHANCK, R. L. *Social psychology.* New York: Wiley, 1938.
14. LEVITT, E. E., & MCCANDLESS, B. R. Some relationships among tests pertaining to the authoritarian personality syndrome in children. *Amer. Psychologist,* 1954, 9, 417. (Abstract)
15. MCCANDLESS, B. R., & HOLLOWAY, H. D. Race prejudice and intolerance of ambiguity in children. *J. abnorm. soc. Psychol.,* 1955, 51, 692-693.
16. MURPHY, G., & MURPHY, LOIS B. *Experimental social psychology.* (Rev. Ed.) New York: Harper, 1937.
17. MUSSEN, P. H. The reliability and validity of the Horowitz Faces Test. *J. abnorm. soc. Psychol.,* 1950, 45, 504-506.
18. MUSSEN, P. H. Some personality and social factors related to changes in children's attitudes toward Negroes. *J. abnorm. soc. Psychol.,* 1950, 45, 423-441.
19. RADKE, MARIAN Y., & SUTHERLAND, JEAN. Children's concepts and attitudes about minority and majority American groups. *J. educ. Psychol.,* 1949, 40, 448-468.
20. ROSENBLITH, JUDITH A. A replication of "some roots of prejudice." *J. abnorm. soc. Psychol.,* 1949, 44, 470-489.

21. Saenger, G. *Social psychology of prejudice.* New York: Harper, 1953.

22. Sherif, H. *An outline of social psychology.* New York: Harper, 1948.

23. Shils, E. A. Authoritarianism: "Right" and "Left." In R. Christie & Marie Jahoda (Eds.), *Studies in the scope and method of the authoritarian personality.* Glencoe, Ill.: Free Press, 1954. Pp. 24-49.

24. Sims, V. M., & Patrick, J. R. Attitude toward the Negro of Northern and Southern college students. *J. soc. Psychol.,* 1936, 7, 192-304.

25. Smith, F. T. An experiment in modifying attitudes toward the Negro. *Teach. Coll. Contr. Educ.,* 1943, No. 887.

26. Snedecor, G. W. *Statistical methods.* (4th Ed.) Ames, Iowa: State College Press, 1953.

27. Springer, Doris V. Awareness of racial differences among preschool children in Hawaii. *Genet. Psychol. Monogr.,* 1950, 41, 215-270.

28. Trager, Helen G., & Radke, Marian Y. *They learn what they live.* New York: Harper, 1952.

29. Vaughn, W. I. *Social psychology.* New York: Odyssey Press, 1948.

THE EFFECT OF CONCEPT KNOWLEDGE ON DISCRIMINATION LEARNING[1]

WINIFRED O. SHEPARD and MAURICE SCHAEFFER[2]
Iowa Child Welfare Research Station

The purpose of the present paper is to describe a group technique for the experimental study of learning processes in children and to report on a study which employed the technique. The study was concerned with the effect of knowledge of a concept name (triangle) on the speed of learning a discrimination problem in which the correct response was always an instance of that concept. The prediction was made that *S*s who knew the concept name would perform better than *S*s who did not.

SUBJECTS

One hundred and seventy-eight sixth grade pupils from six classes in the Cedar Rapids, Iowa, public school system were subjects in the experiment. The classes contained 22, 31, 28, 35, 31 and 31 pupils respectively.

MATERIALS

Twenty stimuli, five instances of each of four classes of geometric figures, were used. There were five triangles, five quadrilaterals, five pentagons and five nonsense figures. These 20 different figures were arranged, as described below, in mimeographed booklets. Each booklet contained 25 pages which constituted the 25 trials of the learning task.

Each page was divided into four 2 × 2 inch compartments labeled A, B, C and D respectively. One triangle, one quadrilateral, one pentagon and one nonsense figure were printed on each page. Thus, each of the 20 geometric figures appeared in the booklet five times. The order of appear-

---

[1] This investigation was supported by a research grant (MH-658C) from the National Institutes of Health, U. S. Public Health Service.

[2] The authors wish to express their thanks to Mr. H. C. Reid, Director of Curriculum and Instruction of the Cedar Rapids Public Schools, and to the teachers and pupils who cooperated in this investigation.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

ance of stimuli was carefully controlled so that there was no tendency for any figure or any class of figures to be systematically associated with a particular position on the pages.

In addition, a concept name test was prepared. This was a mimeographed sheet containing 18 different geometric figures, five of which were triangles.

Procedure

In each class, *E* passed out the booklets and then gave, in essence, the following instructions. "On each page of the booklet you have been given, you will find four figures. Only one of the figures on each page is the correct one. At first you will just be guessing, but after a while you will be able to tell each time which is the correct one. When I say 'ready,' turn to page one and mark the figure you think is correct. When you have done so, raise your hand. When everyone has raised his hand, I will call out the letter of the compartment that contained the correct figure. Then you will turn to page two and do the same thing." This procedure was gone through for 25 trials. The correct choice on each trial was the triangle. As a result of preliminary experimentation with this population, it was found possible to prevent *S*s marking the pages after the correct answer had been called out merely by having *E* walk around the room randomly checking booklets on each trial. Larger classes might require additional proctoring.

Following the learning task, the booklets were collected and the concept name test was passed out. The instructions were simply to write the name of each figure in the space underneath it.

Results

The first step in the analysis was to separate the *S*s into those who knew and did not know the concept name "triangle." Arbitrarily, *S*s who wrote the word "triangle" under each of the five triangles that appeared on the tests were considered to know the concept. *S*s who did not use the word at all or who used it for less than all of the instances were considered not to know the concept. According to this criterion, 58 *S*s knew the concept and 120 did not.

Learning was defined as the occurrence of seven successive correct responses. The probability of seven or more successive correct responses in 25 trials in a situation where there are four possible choices on each trial is .0009 (3). Fifty-four or 93 per cent of those who knew it succeeded in learning the discrimination task. Eighty-seven or 73 per cent of those who did not know it learned. The probability of occurrence of this difference in proportion is less than .001.

In any situation in which *S*s are arbitrarily broken down into two or more groups on the basis of one variable, it is wise to check on the possibility that another variable may be systematically associated with the breakdown and be itself responsible for the observed difference in criterion scores. Accordingly *Ss*' scores on the Otis Self Administering Test of Mental Ability were studied. A two factor analysis of variance (6) of these IQ scores was carried out in which knowledge of the concept constituted one factor and classes the other. The results, corrected for disproportionality (10), showed that neither the concept knowledge factor nor the interaction of that factor with classes approached significance, although there was a significant difference ($.025 > p > .010$) in IQ among the classes. In view of these results, the hypothesis that knowledge of the concept name per se was responsible for the observed difference in criterion scores remains tenable.

One further analysis was carried out to see if still another response measure would reveal the superiority of those who knew the concept over those who did not and to check on the possibility of an interaction between concept knowledge and IQ. This was a 2 × 4 treatments by levels analysis of variance (6) in which knowledge of the concept was the treatment variable and IQ the levels variable, with total number of errors on the learning task as the criterion measure. Four IQ levels were used: 90 and below, 91–105, 106–115, and 116 or over. Due to extreme disproportionality, it was decided to eliminate randomly 16 *Ss*. Hence this analysis, which appears in Table 1, is based on 162 *Ss*. From the table, it can be seen that the treatment variable was highly significant, while neither IQ nor the interaction of IQ with concept knowledge was significant.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF ERROR SCORES

Source	*df*	*Mean Square*	*F*
Concept Knowledge	1	226.67	8.64*
IQ Level	3	30.93	1.18
Concept × IQ	3	18.56	
Within cells	154	26.23	

* $p < .005$.

DISCUSSION

These results can be explained in terms of the notion of mediated generalization based on response-produced stimulation. This is a principle that has been used to account for the fact that *Ss* who have learned one common response to a group of stimuli subsequently have increased tendencies to generalize other responses among them. The theoretical *S–R* paradigm for this phenomenon appears in Figure 1.

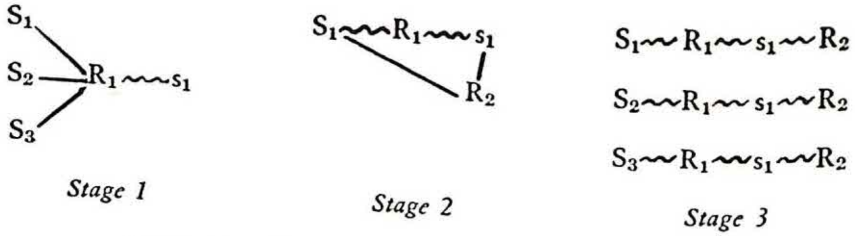

S_1
S_2
S_3
$R_1 \sim s_1$
Stage 1
$S_1 \sim R_1 \sim s_1$
R_2
Stage 2
$S_1 \sim R_1 \sim s_1 \sim R_2$
$S_2 \sim R_1 \sim s_1 \sim R_2$
$S_3 \sim R_1 \sim s_1 \sim R_2$
Stage 3

FIGURE 1—S-R paradigm for mediated generalization

S_1, S_2 and S_3 represent three distinctive stimuli. In Stage 1, *S*s are taught to make response R_1 to each of them. The occurrence of R_1 produces proprioceptive stimulation, s_1, which is assumed to have all the functional properties of any external stimulus. In Stage 2, *S*s are taught a second response, R_2 to S_1. By virtue of the training in Stage 1, S_1 is already known to elicit R_1 with its attendant s_1. During this stage, a connection is established between s_1 and R_2. Now when S_2 and S_3 are presented during Stage 3, R_2 is observed to occur although it had never been specifically conditioned to these stimuli. It occurs because both S_2 and S_3 elicit R_1 and hence s_1, and s_1 elicits R_2. It can also be seen that the Stage 3 stimulus complexes, S_1-s_1, S_2-s_1 and S_3-s_1 are more similar to each other than are S_1, S_2 and S_3 alone which appear in Stage 1. Hence generalization tendencies are higher in Stage 3.

In the paradigm, R_1 is referred to as the mediating response. Theoretically, any common response learned to a set of stimuli could serve as a mediating response. Birge (1) and Jeffrey (5) have, in fact, shown that both verbal and motor responses can be mediators.

In the present study the verbal response "triangle" served as a mediating response. Presumably, prior to the experiment some *S*s had learned to associate this response with every instance of a closed three-sided figure. During the learning task this verbal response produced proprioceptive stimulation to which the marking response became conditioned. After the first correct marking response was made an association was begun between the stimulation produced by the word "triangle" and the response of marking the figure which elicited that name. This association became rapidly stronger with each succeeding correct response. As soon as these *S*s turned a page, one stimulus on that page elicited the word "triangle" and hence the marking response. The problem for these *S*s had been learned.

On the other hand, *S*s who did not use the word "triangle" at all, or who used it for only some types of triangle, were faced with a different situation. They could learn the task in one of two ways. One way would be to simply learn the marking response separately to five different stimuli—the five triangles which appeared throughout the booklet. [Gellerman (2) showed that both monkeys and very young preverbal children could learn the concept triangle in this way.] Or, these *Ss* could formulate for themselves, during the learning task, some common response such as "three-sided" which could serve to mediate correctly the generalization of the marking response. Either of these methods would take longer than the immediate application of an already available mediating response. Hence, in comparing the performance of *Ss* who did and did not know the verbal response "triangle," a larger proportion of the latter did not learn at all in the time allowed and, in general, they made more errors.

It should be noted that these findings are consistent with those of other studies, Heidbreder (4), Long and Welch (7, 8), Roberts (9), Snygg (11) and Welch and Long (12) in which it was found that children who can verbalize the basis of learning show superior performance to those who cannot, even in cases where nonverbalizers eventually do learn.

A final word remains to be said about the particular technique of gathering data employed. It is, of course, possible with this group technique to gather in a few hours the same amount of data that it takes months to gather when working with *Ss*, particularly children, on an individual basis. It is furthermore felt that the technique is amenable to the study of many discrimination learning variables. For instance, the similarity of the stimuli to each other, their number and complexity could be systematically varied. Also, motivational variables could be manipulated through the use of printed instructions for *Ss* old enough to read.

SUMMARY

A group technique was used in studying the effect of concept name knowledge on learning in sixth grade children. The task involved choosing the triangle on every trial from among four geometric figures. There were five different instances of triangles throughout the task. Children who demonstrated on a concept name test that they could consistently and correctly use the word "triangle" were found to perform better on the learning task than children who could not. The results were discussed in terms of the mediated generalization hypothesis.

REFERENCES

1. BIRGE, JANET. The role of verbal responses in transfer. Unpublished doctor's dissertation, Yale Univer., 1941.
2. GELLERMAN, L. W. Form discrimination in chimpanzees and two year old children. *J. genet. Psychol.*, 1931, 39, 50-72.
3. GRANT, DAVID A. Additional tables of the probability of "runs" of correct responses in learning and problem solving. *Psychol. Bull.*, 1947, 44, 276-279.

4. Heidbreder, E. Reasons used in problem solving. *J. exp. Psychol.,* 1927, 10, 397-414.

5. Jeffrey, W. E. The effects of verbal and nonverbal responses in mediating an instrumental act. *J. exp. Psychol.,* 1953, 45, 327-333.

6. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

7. Long, L., & Welch, L. Reasoning ability in young children. *J. Psychol.,* 1941, 12, 21-44.

8. Long, L., & Welch, L. Influence of levels of abstraction on reasoning ability. *J. Psychol.,* 1942, 13, 41-59.

9. Roberts, K. E. Ability of preschool children to solve problems in which a principle of relationship is kept constant. *J. genet. Psychol.,* 1932, 40, 118-133.

10. Snedecor, G. W. *Statistical methods.* Ames: Iowa State College Press, 1946.

11. Snygg, D. Relative difficulty of mechanically equivalent tasks. I. Human learning. *J. genet. Psychol.,* 1935, 47, 299-320.

12. Welch, L., & Long, L. Comparison of reasoning ability in two groups. *J. genet. Psychol.,* 1943, 62, 63-76.

Supplement to

CHILD DEVELOPMENT

Volume 27
June, 1956
No. 2

CHICAGO CONFERENCE ON CHILD DEVELOPMENT

1955

Contents

Child Development: An Historical Perspective 181
John E. Anderson ....

Some Problems of Personality Development Research 197
Vaughn J. Crandall ....

A Program of Research on the Dimensions and Antecedents of
Emotional Dependence 205
Jacob L. Gewirtz ....

Some Research Problems for Developmental Psychology 223
Eugene S. Gollin ....

The Replacement Problem in Child Development 237
Harold E. Jones ....

The Need for Conceptualization in Research on Child Development 241
Irving E. Sigel ....

Experiments with Children on the Hypotheses of Acquired
Distinctiveness and Equivalence of Cues 253
Charles C. Spiker ....

Psychological Development in Midwest 265
Herbert F. Wright ....

CHILD DEVELOPMENT: AN HISTORICAL PERSPECTIVE

JOHN E. ANDERSON
Institute of Child Welfare
University of Minnesota

As with development, history may be approached from several points of view, each of which results in a different treatment. First, there is a straightforward chronicle of events, a descriptive account of what happens, which gives minimal determination of relations other than those indicated by the time order itself. As with longitudinal studies of development, mountains of records become available from which selection occurs in both recording and retention.

A second approach is in terms of men, since particular persons acting as leaders express group tendencies and gather followers who thenceforth affect the stream of relations. Analysis in terms of leaders however identifies the social process with a few individuals. It is the history of the priests rather than that of the people. Particularly is this the case in the history of science, which, in its very essence, depends upon many intellectual workers who pursue similar lines of inquiry.

Third, there is an analysis of the trends within historical time based upon antecedent-consequent relations and the interconnections of events. This reveals the essential logic of the historical process and outlines the emergence of ideas and of intellectual movements. But this approach is, however, open to great error since a complex manifold is approached from the point of view of a theory in terms of which the material in the manifold is restructured. Very often there is no possibility of validation since history as an open end system can never be reconstituted. This problem also faces us in the study of human development.

Moreover, the symbolic system into which history must be put is quite as inadequate as is our symbolic system for representing the developmental process. Simple relations between two events, persons, variables, sequences are well handled; the complex of relations and the relations between relations make up a manifold that can probably never be represented as a totality. Our problem is complicated further by the fact that the system

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

is dead and gone; we push the dry bones of our records around in the hope that some glimmer of the life that was will emerge.

PRECURSORS

Before discussing child development which begins as an entity in the early twenties, we may take a hasty glance at thinking about children in an earlier period. It is difficult for any of us to understand the world of one hundred years ago in which children moved through school without clear recognition of differences in capacity and growth, without understanding of the learning process, and without insight into what underlies good and bad adjustment. It was a world in which the study of children was frowned upon. G. Stanley Hall has told how difficult it was in the middle of the last century to measure the height and weight of school children because parents objected to the invasion of the child's rights. The few persons who kept records of children were scientists who used their own children as subjects. Usually they were satisfied with the production of a baby biography.

Recently I read a text widely used in 1888 for training teachers in what would now be called child development. In it there is no discussion of individual differences or of growth; only three pages are given to what would now be called learning. Memory is discussed in twelve pages without any data. Most of the book is spent on sensation and perception and on material on reasoning which now would be found mainly in texts on logic or philosophy. There are no quantitative materials, no citations of experiments, and no tables or graphs.

In the nineties several intrepid investigators such as Bolton and Gilbert tested reaction time, color discrimination and memory in children from six to sixteen. But unfortunately their tests showed little or no relation to school performance and had no practical results in placement or training. So they remained ahead of their time.

Great strides came from experiments on learning on adults. Bryan and Harter plotted their observations on learning telegraphy and discovered the learning curve. With this demonstration that progress in learning could be plotted and measured, it became clear that methods of improving instruction and learning could be discovered by experiment. This area of research was rapidly extended to children. The pioneer experiment of William James on transfer of training was followed by the important investigations of Thorndike and Woodworth in 1901, which attacked the doctrine of formal discipline and set off a great controversy and much research, most of which was done on children.

Now comes a great nodal point. In 1907, Alfred Binet, impressed by the injustice of the intuitive method of placing children in institutions then in vogue, sought an objective method for assessing the child's level that would be independent of the examiner's judgment. He came upon the method of age scaling and demonstrated that children's performance increases rapidly with age. Soon it was clear that the measurable differences in mental level affected children's school achievement. Although differences were recognized prior to Binet, they were felt to be related to working habits, moral fiber, and willingness; in other words, to character and moral qualities over which the child had voluntary control. And let us not forget that it was a world in which many children were given severe physical punishment because they had what we now recognize as inadequate capacity.

Why is Binet's work so important? Clearly it substitutes for the onus of moral blame a measurable phenomenon within the child's resources, which is to be studied henceforth in its own right. The repercussions were tremendous. Bright children, normal children, handicapped children, retarded children, deafened children, blind children; all types and kinds of children everywhere were measured and studied in the years that followed. But, more than that, Binet made psychologists everywhere so highly aware of the relation of chronological age to mental processes that henceforth almost every investigation on children contains some analysis of the results in terms of chronological or mental age. Moreover, Binet's techniques spread widely and rapidly over the whole world. It is impossible for us, unless we lived through the period, to recapture the enthusiasm, discussion and controversy that the Binet tests started.

Then came Thorndike, who in addition to working out statistical techniques, stimulated vast research on almost everything that could be measured in school children. From this there came critical attitudes toward long-existing practices that resulted in many experiments designed to evaluate the educational process.

Here then are a number of trends in the psychological field which came together: intelligence tests, learning, transfer and interference experiments, and widespread concern with the relation between the products of education and age level. There resulted an atmosphere which made age relations, learning, and individual differences focal points for study.

Then World War I gave an opportunity to apply some of the testing techniques which had grown out of research on children to a young adult population in the military service. Although the tremendous effects of the psychological work of World War I in developing measuring techniques is beyond our scope, we may point out that hundreds of psychologists who were trained in the new measurement techniques were spread out over the entire United States after the war. This very rapid diffusion of knowledge had immediate and great consequences in providing the techniques with which a growing science attacked many human problems including those of children. A third and very striking effect of World War I with its physical, medical and psychological examinations was to present a picture of the young adult American male, which was not very pleasing. A surprising number of young men showed correctable physical defects. Illiteracy, because of inadequate educational facilities in many sections of the country, was clearly demonstrated. Many maladjusted and inadequate persons were discovered. Educational opportunities were clearly below the capacity of many who had been pupils.

THE CHILD DEVELOPMENT MOVEMENT

Hence, the early 1920's were times of great ferment, which extended into medicine and health as well. As the American people set out to remedy the defects in the physique and the training of their young men, new sciences and new arts relating to the physical and mental development of human beings came into being. This was the period when well-baby clinics, preventive pediatrics, school health programs, child guidance clinics, mental hygiene clinics, and demonstrations in public health put in their appearance. And among the new movements was the child development movement, which this time became a reality. Once before, in 1909, under the sponsorship of G. Stanley Hall, an attempt had been made to coordinate the activities of various scientific societies into an interdisciplinary attack upon developmental problems. But the atmosphere then was not right.

Systematic and organized study in the field did not begin until the early twenties when great interest in the growth and development of children appeared. In 1917 the Iowa Child Welfare Research Station opened with Bird T. Baldwin as Director and in 1921 it published a study of the physical growth of children which was followed a year later by a monograph on the mental growth curves of normal and superior children. In 1916 Terman began working with gifted children. In 1919 Watson's *Psychology from the Standpoint of a Behaviorist* with its strong emphasis on the importance of data on infants for the understanding of adult psychology appeared. In the early twenties, Gesell at Yale was observing the behavioral development of infants and young children, while Dearborn began the Harvard Growth Study and Freeman the Chicago Growth Study in each of which large school populations were followed from year to year with physical measurements and mental tests. Much interest in development appeared in different institutions and places. Of importance also were the continuing studies of Thorndike and his students on ability and achievement tests for children of various ages and school levels.

Spelman Funds

Into this picture there came the Laura Spelman Rockefeller Memorial with two able and active persons, Beardsley Ruml and Lawrence K. Frank, who, to meet one of the avowed purposes of the Memorial, were responsible for encouraging research in child development and parent education over the whole United States. Frank in particular played a prominent part in the early organization of the field and maintained high personal interest and enthusiasm for the activities undertaken. Substantial funds were made available and a series of institutes whose primary function was the conduct of research in child development were established in universities. The Child Welfare Institute at Teachers College was established at Columbia in 1924, and the Institutes of Child Welfare at the University of Minnesota and at Toronto in 1925. In 1926 the Yale Psycho-Clinic was financed. In 1927 the Institute of Child Welfare at the University of California was established and substantial funds were made available to the Iowa Child Welfare Research Station. Smaller grants were also given but these were mainly concerned with the education of very young children or with parent education. In order to increase the effectiveness of these grants, funds were made available to the Child Development Committee of the Research Council for fellowships, conferences for stimulating research and for its work as a coordinating agency. The plan of the Laura Spelman Rockefeller Memorial was comprehensive. Put into action, it literally created a field that was nation-wide in its scope and thus advanced work in the area of child study by several decades.

In many of the units established under Spelman Funds, nursery schools were established in order to make young children available for observation and study. In some, funds were available for parent education programs. In several, special facilities or projects concerned with the observation of infant behavior were established. Although the primary emphasis was on the infant and young child, some projects were undertaken at later ages. As the time passed the Institutes grew with their children and more studies were undertaken on older children.

For the infant studies the most widely known center was the Yale Psycho-Clinic where Gesell developed his observational dome and made follow-up studies of many infants. But there was great interest in infant development at other places: Weiss' studies at Ohio State; the Shirley and Boyd study at Minnesota; the McGraw study which made Johnny and Jimmy famous; and the various studies of Bayley *et al.* at California.

Compared with the studies on infants which were confined to a few centers, there were many studies on children from two to five years. A whole scientific literature covering virtually every aspect of the developmental process was created within a few years by the activity of the various Institutes. Into the picture there also came a number of organizations with other sources of funds. One of these was the Merrill-Palmer School at Detroit which began in 1920, but did not become active in research until the mid-twenties. At Cornell a program in Child Development and Family Life began in 1925. The Brush Foundation at Western Reserve University was established in 1927 to carry on longitudinal studies and the Mooseheart Laboratory for Child Research was established in 1930. In 1927 the Child Research Council at Denver began operating. Later it amalgamated

OUTLINE: HISTORY OF CHILD DEVELOPMENT

	Before 1920	1920--29	1930--39	1940--On
University Units	Individuals Darwin 1809 Galton 1822 Preyer 1841 G. S. Hall 1846 Binet 1857 J. M. Baldwin 1861 Thorndike 1874 Stern 1871 Wallin 1876 Koffka 1886 Buhler, C. 1893 Werner 1890 Piaget 1896 Freud 1856 Watson 1878 10 Gatzert Fund became ICW, U Washington (Smith 83) 17 Iowa C W Research Station (Baldwin 75, Sears 08)	24--34 CWI, Teachers College, Columbia (Wooley 74, Jersild 02) 25 ICW, Minnesota (Anderson 93, Scammon 83, Goodenough 86) 25 ICW, Toronto (Blatz 95) 25 CD&FL, Cornell (Waring) 26 Yale Psycho-Clinic (Gesell 80) 27 ICW, California (Jones 94) 27 Child Research Council, Colorado (Washburn 95) 28 Brush Foundation, W.R.U. (Todd 85)	30 Research in Child Dev., Michigan (Olson 99) Committee Human Dev., Chicago	48 Child Study Dept., Yale
College Units		29 Fels Institute, Antioch (Sontag 01)	35 Char. Research Proj., Union (Ligon 97)	
Clinics	96 Psychological Clinic, University of Pennsylvania (Witmer 67)			
Independent Units	08 Vineland Research Laboratory (Goddard 66, Porteus 83, Doll 89) 08 Elizabeth McCormick Memorial Fund	20 Merrill-Palmer School (White 80, McGinnis 93)	30 Mooseheart Laboratory for Child Research (Reymert 83)	
Longitudinal Projects		22 Stanford Gifted Children (Terman 77) Harvard Growth Study (Dearborn 78) Chicago Growth Study (Freeman 80)	30 Neuromuscular Dev. of Infants and Johnny and Jimmy (McGraw 98)	
Handbook or Manual			31 (Murchison 87)	46 (Carmichael 98)
Foundations	18--38 Laura Spelman Rockefeller Memorial (Ruml 94, Frank 90) 18 Commonwealth Fund	24 Child Guidance Clinics (Stevenson 92)		
U.S. Government	12 United States Children's Bureau White House Conferences 1909 .... 1919	1929		1940
Voluntary Lay Organizations	88 Child Study Association of New York (Gruenberg 81) 97 National Congress of Parents and Teachers 98 Child Study Program, American Association of University Women	25 Association for Family Living, Chicago (Lawton)		
Professional Organizations	1909 Hall's abortive attempt	20 Committee on Child Development, National Research Council (Woodworth 69) 26 National Association Nursery Education 26--41 National Council of Parent Education	33 Society for Research in Child Development	46 Division of Childhood and Adolescence, A.P.A.

with the University of Colorado and started a series of extensive longitudinal studies. In 1929 the Fels Institute at Antioch began its extensive program covering many aspects of child development.

In order to coordinate the efforts of various scientists and to further interchange of information about research techniques, the National Research Council revitalized its committee on Child Development in 1925. In the course of the next few years this committee held a series of conferences in which representatives from various sections of the United States and from various disciplines participated. The committee had an active secretariat with funds from the Laura Spelman Memorial. It also became an agency through which substantial sums were distributed to fellows and scholars in child development and in parent education.

Out of the activities of this committee came the Society for Research in Child Development, which was organized in 1933, and which has continued up to the present time with its various publications.

In discussing the development of the field some attention must go to the White House Conferences. The first of these called in 1909 by President Theodore Roosevelt concerned itself mainly with what the national government should do about children. Out of it came the U. S. Children's Bureau, which has been active ever since. The Conference in 1919 dramatically brought to the fore the physical and mental deficiencies revealed by the draft in World War I. The Conference in 1929 pooled substantial amounts of research from all the agencies in the United States (including much data from the Institutes), conducted various surveys on its own, and issued a substantial series of publications which are still significant. The Conferences held in 1940 and 1950, however, were much more concerned with the practical and social aspects of the problems presented by children rather than with research findings.

As an aid to visualizing this chronology there is attached an outline which locates in the time sequence various (not necessarily all) units beginning with the year 1920. In the upper left corner are the names of persons who were prominent in the field prior to or parallel to the formal units. Wherever a person is identified with one of the units his name appears immediately after along with the date of his birth.

MEN

It is not my purpose to single out particular persons for comment but to generalize with reference to types of contributions made. Before 1920 the men who were interested in research on children were scientists with many widespread interests. Some, such as Galton and Darwin, were mainly interested in the whole developmental process from a biological background; some such as Preyer were physiologists who turned to embryology and the early stages of life; some such as Romanes and J. Mark Baldwin had philosophical as well as biological interests and sought answers to phylogeny in ontogeny. There were no specialists in the modern sense, even though G. Stanley Hall probably comes nearest. Yet his interests were very broad. For Binet, whose influence in the public mind is closely identified with the single outcome of intelligence testing, the development of the scale was an incident in a career of many and varied contributions to psychology.

Moving into the twenties we find child development selecting its personnel from many intellectual and scientific sources. With the establishment of the Institutes what amounted to a profession was established. Even yet, however, we are in the throes of the ongoing developmental process with much controversy whether or not we are multidiscipline, interdiscipline, single discipline or an offshoot of a particular discipline. But nevertheless a considerable group of dedicated persons identify with child development as a field.

In talking with a friendly historian he made the point that in writing history one must look out for persons who exercise great effect upon a field without leaving behind the permanent records which become the material for history. Often devious searches and odd clues must be used to locate them. In the living system they are powerful determinants; in the dead system they leave few traces. We have such a person. In the McLean article on the history of Child Development (5), R. S. Woodworth is barely mentioned as a member of the Committee on Child Development. No account is given of his long chairmanship of the committee, his efforts in organizing the Society for Research in Child Development, his great services in securing recognition and status for child development in the National Research Council and the Social Research Council, and his active participation in conferences, committees, meetings and in the development of publication policies. He is the strong silent man of our history, a man famous enough outside our field, who was for us over many years a real father figure.

TRENDS IN CONTENT AND IDEAS

Growth of the Scientific Literature

One approach that can be made in evaluating the changes in time within our field is to make some counts of the scientific literature in the field. This is particularly important in view of the fact that there are the Spengler's among us such as Dennis and Barker who give the impression that we are going downhill in our knowledge of development and are farther from our goals than we were in 1925. To approach this problem I made the three following analyses.

Textbooks. The first book with the title *Child Development* that I have found is that of King (3), an associate of John Dewey, published in 1903. When one compares this book with a modern text the differences in 53 years are impressive. King's book is quite general and philosophical and, except for an area on the activities and interests of children which refers to many of the studies of G. Stanley Hall and his students, there is a paucity of references to specific data. Compared to Thompson's *Child Psychology* (1952) with its 667 pages, its 225,000 words of text, its 1177 references, its 150 figures and its 41 tables; King has 265 pages, 45,000 words of text, 106 references, 4 figures on physical growth (as a sort of appendix) and no tables.

Handbook and Manual. A check of the *Handbook of Child Psychology* (6) and its successor the *Manual of Child Psychology* (1) makes obvious the increase in the accumulated material in the field as shown in Table 1 which gives the number of words for the book as a whole and for the only two chapters (Anderson, McCarthy) which have covered the same content in all four editions. Obviously such figures only approximate the trend since editorial considerations determine in some degree the amount of material presented.

TABLE 1

WORDS* IN *HANDBOOK OF CHILD PSYCHOLOGY* AND *MANUAL OF CHILD PSYCHOLOGY*

	Entire Book	*Anderson Methods*	*McCarthy Language*
Handbook 1st Ed. 1931 . . .	275,000	11,000	14,500
Handbook 2nd Ed. 1933 . . .	405,000	13,500	18,000
Manual 1st Ed. 1946 . . .	525,000	23,000	56,000
Manual 2nd Ed. 1954 . . .	650,000	32,500	73,000

* Exclusive of bibliographies.

References in Psychological Abstracts. Recently I have had the references entered in the yearly indexes of *Psychological Abstracts* counted under the headings "Child," "Infant," "Adolescence," and "Family" from the first volume in 1927 up to the present. The results are presented in Table 2 by seven year periods.

Far from showing a decline in interest in children or research upon them, these figures show an increase in the absolute number of references and a fairly constant percentage relation of about 7 per cent which indicates that publication in the child field varies with that in the total field. "Infant," "Adolescence," and "Family" have a small number of references with some increase in absolute numbers and fairly constant percentages.

While these different analyses concern only the quantity and not the quality of our intellectual and scholarly production, they certainly do not indicate that we are worse off than we were in an earlier period. While the Golden Age in which the Institutes were being established and which is

TABLE 2

MEAN NUMBER OF INDEX REFERENCES AND PERCENTAGE OF TOTAL PER YEAR FOR VARIOUS TOPICS IN *PSYCHOLOGICAL ABSTRACTS* FOR SEVEN YEAR PERIODS FROM 1927 TO 1954

	Child		*Infant*		*Adolescence*		*Family*		*Total Abstracts*
Period	Mean per yr.	%	Mean per yr.	%	Mean per yr.	%	Mean per yr.	%	Mean per yr.
1927–1933	325	6.7	8	0.2	38	0.8	29	0.6	4704
1934–1940	437	7.0	45	0.7	53	0.8	30	0.5	6270
1941–1947	386	8.5	27	0.6	44	1.0	30	0.6	4555
1948–1953	501	6.9	36	0.5	72	1.0	62	0.9	7362
1927–1953	412	7.2	29	0.5	52	0.9	38	0.7	5723

reflected in the publications for the late twenties and early thirties is over, it is my opinion that there has been an increase in interest in research on children throughout the entire field of psychology (the field with which I am most familiar) and that the Institutes are no longer quite as much in the center of the picture.

Approaches to the Study of Growth and Development

In considering the content of child development we need some conception of the nature of the growing human person or a theoretical model against which to orient our studies. The growing person is an ongoing manifold moving through time with a multidimensional head much as a tunnel-making machine moves under a river. Viewed from the moving front or the present moment of time this consists of many irregularly shaped structures of various sizes with a complex series of interrelations. At the boundaries there are complex relations to the external world which involve a continuous intake and outgo on both a physiological and a psychological level, an interchange which maintains a dynamic or ongoing equilibrium.

Viewed as a whole the system is asymmetrical and open-end. Its basic organization is hierarchical in the sense that as it moves forward in time new properties and qualities emerge at successive levels of integration. These need study and analyses in and for themselves. An analysis at one level, however adequate, cannot explain the properties of a higher level.

In other words we cannot solve the problems of children's language by analyzing their pronunciation of single letters or syllables, but must study his behavior for words and sentences. The social behavior of the child must be studied in its own right by its own methods since there are limits to the extent to which it can be predicted from knowledge of the individual children in the group.

At once this model denies the possibilities of explaining human behavior on the basis of physical models and makes the psychological model analagous to those of theoretical biology. In exploring the possibilities of this model we are concerned with the study of its total shape or form as well as the shape, form, and interrelations of the parts. Moreover, analysis of time relations and of transformations in time become crucial.

Many of the apparent contradictions in behavior and development theory result from the fact that we study different aspects of this exceedingly complex manifold as it undergoes transformation in time. To the complexities of the unfolding system are added the inadequacies of our symbolic system which, whether verbal, pictorial or mathematical, has distinct limitations (9).

Using our model as a base we may have a look at the accumulated scientific literature in our field which we can divide into eight broad classes, each of which could be treated in considerable detail.

1. *Space-time form studies.* These are concerned with the properties and interrelations of systems within the whole model as it evolves in time. After the comparable biological science which studies the transformations in physical growth, this area is a kind of psychological morphogenesis. Either successive cross section slices or a combination of many longitudinal slices may be used to create the picture. Although such studies result in a wealth of information about the parts of the model, because of their rarity emphasis could well go not so much to the slices or factors as such as to the interrelations between variables and the emergence of higher levels of behavior. This is the most involved type of study and, except for the longitudinal studies at California, we have hardly made a dent in this area.

2. *Longitudinal slices.* There are a number of studies in which some particular or several aspects of development are followed in the same subjects over a period of time. The distinction between these slices and the studies under the previous heading is one of extent. Related to both are the follow-up studies, now being set up in many places in which records obtained much earlier are used as a base to study continuities or discontinuities. The Terman studies of intelligence are an example in point.

3. *Cross section studies at successive levels.* These analyze segments or relations over substantial periods of time by selecting particular aspects of development at successive age levels on comparable samples of the population in order to plot or lay out the course of development. These studies become concerned with the constancy or change problem, and with the analysis of the underlying factors which produce constancy or change. Problems of definition and classification of the ongoing phenomena at successive age levels also arise. There is a great bulk of studies of this type in many areas. At present we can write a respectable picture based upon substantial and reasonably complete research of the development of the child from birth to maturity in many psychological areas, particularly intel-

ligence and language. There are, however, some gaps in our knowledge of perceptual, motor, concept, personality and social development.

4. *Cross section studies at one level.* Many persons are interested in children at a particular level. Some are interested in infants, some in young children, some in adolescents, etc., and think only in terms of the particular level. In the design of experiments this approach leads to cross section slices taken on various samples without forward or backward reference in the developmental sequence. There are literally thousands of such studies essentially at a descriptive level which are difficult to assess or to fit into a comprehensive picture of development. It seems to me that the day of such studies is pretty well over, and that we should urge scientists who work with children to study at least three separate age levels in order to determine the relation of the phenomena in question to age. If the primary concern of the developmentalist is with changes in time, chronological age or some derivative of it must become the independent variable in our studies.

5. *Short-period function modifications.* These introduce specific stimulation compressed into a relatively short time on a sample of children. They are concerned with the mechanics of stimulation and response, or the manner in which patterns of behavior develop. One method of instruction is compared with another or a learning or problem solving situation is set up. In terms of design these studies select groups of children at some level, introduce stimulation for a time and then measure response. Learning curves are plotted and analysis is made. The experimenter loses interest in the child as soon as he reaches one or ten perfect trials or attains a defined level of proficiency and is not interested in what the child does with what is learned, nor how the content is incorporated into the total system of behavior. Rarely are such problems carried out on children at several age levels, as the investigator is only incidentally concerned with the developmental ranges through which the principles derived operate. Although the literature on perception and learning in children is substantial, it is much more concerned with simple than with complex processes.

6. *Genetic determinism.* Many studies are concerned with the unfolding or the relations of hereditary factors. Resemblance coefficients between parents and children and between siblings are commonly used. We need more and better studies of the contribution of genetic variance to the total variance in behavior, and particularly studies in which children are examined at the same ages as their parents were, in order to minimize differences due to developmental level.

7. *Environmental impingement and determinism.* These study the positive effects upon the whole and its parts, of the environment as a whole or in part along with the negative effects of deprivations of various types. Here again we move into a vast array of studies many of which are of very limited scope and offer little or no possibility of effective integration into the literature. Many are poorly designed, done on very limited age samples,

fail to isolate significant factors, and, by making superficial analyses of the relations involved, often create artifacts. Many fail to test for the permanence of the effects. In the deprivation studies which are so popular now, there is often lack of good controls, failure to recognize the basic capacity of the organism for self repair when returned to a good environment, and a general lack of sophistication about scientific method. But studies of the responses of children in their habitat, of the Barker and Wright type, offer new methods which can be widely used and suggest an entropy theory of stimulation rather than the common nickel-in-the-slot one.

8. *Construct determinism.* These studies reify a quality or attribute of the system and study the outcomes of the construct so derived. One phenomenon, usually motivational, is generalized and its effects studied. Thus we have frustration, anxiety, rejection reified into factors which thereafter determine behavior. Similarly we reify the "only child" status and ordinal position into determiners. The reified characteristic becomes a sort of product which once deposited in the ongoing system operates virtually indefinitely without modification. Here we need to find the iterative mechanisms that keep constructs, once we are sure of their validity, alive and active in the developing system. Many of the studies in this area are made on very small samples of cases at limited age ranges, but with wide generalizations to populations of all ages and types. In addition to substantial samples emphasis should go to observations on at least three age levels in order to discover what happens with age.

Areas of Study

As Woodger (9) has pointed out, the first approach to a growing organism involves cataloging and describing the many structures, functions and behaviors of the organism as it moves through time. This cataloging function which is typical of the early stages of life sciences is essential in order to locate the problems and to develop the hypotheses that make more complex generalizations possible. In the period from 1920 to 1930 workers in child development and related fields were mapping out details of physical and mental growth and particularly those from birth to six years. By fostering infant laboratories and nursery schools within the research organizations, Spelman financing directed the attention of many workers to the infant and the young child. There was much concern with the standardization of measuring instruments, such as intelligence tests for infants and young children, and much interest in language, problem solving and social behavior. Perhaps the most original developments of this era lay in the field of social behavior as the extraordinarily large number of investigations summarized by Murphy, Murphy and Newcomb (7) indicates. The sociometric techniques were also emerging. There were scattered investigations on the learning of young children.

At the White House Conference in 1929-30 a committee selected a stratified sample of young children over the whole United States for study of the relation of socio-economic status to standards, facilities and methods of care. There followed many studies of effect of social class and social status upon the development of the child with very clear demonstrations that socio-economic or cultural status is a significant determiner of variance in the goals and ideals of the parents, the methods of child training and the facilities made available to them for learning and adjustment.

Another trend resulted from the studies in the late twenties on diet selection which led to an almost complete reversal of practices with regard to schedules in the training of infants and to the movement toward self-demand rather than forcing in the socialization of the infant and young child.

Lewin and his students in the late thirties performed the experiments on autocratic and democratic atmospheres which resulted in substantial research and new conceptions of the relation of the child to its environment. The principles were later extended upward into adult behavior and revitalized areas in social psychology. In our field came the studies of home and family atmospheres with much improvement in our descriptions and analysis of the family environment and of parental behavior. Meanwhile, however, the cataloging function continued and the first results of the longitudinal studies began to appear in substantial degree. In the early 1940's psychoanalytic concepts, particularly relating to the effect of deprivation and institutionalization or the relation of early handling of the child and subsequent development, began to appear and much concern, both experimentally and theoretically, for the last decade and a half has centered in that area.

Then came World War II which temporarily diverted many of the workers in the field directly to war activities and others to such practical matters as the care of children during the emergency and technical assistance in the war or defense efforts. The organizations devoted to research on children kept going on in one way or another, but with reduced loads and sometimes with change in emphasis in their research program. So great were the disturbances that it took five or more years to return to an even keel.

To some degree just prior to the War and strikingly thereafter came the marked developments in projective techniques which opened new areas of personality measurement for both children and adults. These are still centers of controversy as well as of active research. After the war there appeared the habitat studies and techniques of Barker and Wright with their complete description of the child-environment interaction and the conceptualization of methods for analyzing the environment.

Recently there is a striking new development represented by the follow-up studies based either upon a longitudinal sample earlier followed for many years or the follow-up as adults of the many children on whom the Institutes obtained records in the late twenties and thirties. In some degree the research of the Institutes has grown older with their children. Some effort is going into late adolescence and early adult life. Whether with these developments as much research energy can go to young children as in the early days remains to be seen.

CONCLUSION

In my summary of the developing content within the field I have used three approaches: first, quantitative estimates of the growth of the scientific literature; second, an analysis of our different approaches to the model of the ongoing manifold changing in time which we know as the growing human being; and, third, a brief summary of developments in content and methods by areas in the past three decades. None of these gives us any reason for discouragement or apology. A scientific literature of wide scope has been created where there was almost none before.

In large part this is the product of persons and institutions who became interested in human development in the twenties and who with their students became the professionals in the field. But also because development has acquired great significance in theories of behavior and adjustment, much work has been undertaken outside our channels. The briefest search of the extant literature will show such interests widely scattered among psychologists, sociologists and anthropologists. And quite recently a surprising amount of interest in the young of the species has appeared among the students of animal behavior.

But even more impressive than the extent of the literature are the new methods of approaching the behavior of the child, of analyzing the stimulation of the environment, and of studying the modification of behavior. These give us a surprisingly rich armament with which to move forward. And, as with any science in a period of rapid progress, new problems continue to emerge.

Finally I feel that we are essentially a biological science, studying an open-end living system of extraordinary complexity. We cannot hope for the extreme condensation of modern physics; the equation of a living organism, if such could be written, would fill many blackboards. We have done much cataloguing and describing and must continue to do so. But this is not enough; we must move on to first order generalizations that are close to our data and then on to those of the second, third and *n* orders. Whenever we can substitute experimentation for description we will move ahead. But let us not assume that a theoretical model erected for a limited portion of the relational system is the equivalent of the model for the whole. We need better methodology; we need better theory; we need more adequate symbolic systems if we are to picture and conceptualize what has been often described as the most interesting phenomenon of life: the fertilized cell becoming the complex manifold that we know as the mature human being.

REFERENCES

1. CARMICHAEL, L. (Ed.) *Manual of child psychology.* New York: Wiley, 1946 (1st Ed., 1068 p.), 1954 (2nd Ed., 1295 p.).
2. DENNIS, W. Historical beginnings of child psychology. *Psychol. Bull.,* 1949, 46, 224-235.
3. KING, I. *The psychology of child development.* Chicago: Univer. of Chicago Press, 1903.
4. KIRKPATRICK, E. A. Child study. *Pop. Sci. Mon.,* 1910, Nov.
5. MCLEAN, DOROTHY. Child development: A generation of research. *Child Develpm.,* 1954, 25, 3-8.
6. MURCHISON, C. (Ed.) *Handbook of child psychology.* Worcester, Mass.: Clark Univer. Press, 1931 (1st Ed., 711 p.), 1933 (2nd Rev. Ed., 956 p.).
7. MURPHY, G., MURPHY, LOIS B., & NEWCOMB, T. M. *Experimental social psychology.* New York: Harper, 1931 (Rev. Ed., 1937).
8. THOMPSON, G. *Child psychology.* Boston: Houghton Mifflin, 1952.
9. WOODGER, J. H. *Biology and language.* Cambridge: Cambridge Univer. Press, 1952.
10. *Psychological Abstracts,* Vols. 1-29. Washington, D.C.: Amer. Psychol. Ass., 1927-1954.

SOME PROBLEMS OF PERSONALITY DEVELOPMENT RESEARCH

VAUGHN J. CRANDALL
Fels Research Institute for the Study of Human Development

The content and methodology of personality development research, like those of the total field of child development research, vary tremendously. Past personality development studies have ranged from empirical normative surveys to theory-oriented hypothetical-deductive ones, and from naturalistic observational investigations to controlled laboratory experiments. No attempt will be made in the present paper to survey all, or even most, of the current theoretical and methodological problems in this area of child development research. Rather, a limited number of these problems will be selected for discussion, with special emphasis on their implications for future research.

A number of personality development studies in the recent past have employed projective techniques as research tools. While some of these techniques have been very useful for studying certain problems, projective techniques have often, unfortunately, been employed in an uncritical fashion. Too frequently the tail has wagged the dog, with research problems being determined by available projective methods rather than by specific theoretical questions. It is the belief of the writer that this state of affairs can be, and probably will be, eradicated as research in this area becomes less empirical and more theoretically oriented. Ideally, in theory-oriented research, hypotheses under consideration should dictate behaviors to be studied and thus, indirectly, determine the choice of assessment techniques to be used. Existing projective techniques may or may not be appropriate for such investigations. In sheer number there is no shortage of potentially useful projective tests; the growth of these assessment methods has been almost weed-like. However, since many of these techniques were originally developed with little regard for existing personality theories, few may be found useful in future personality development research in their present form. New techniques will have to be developed and/or old ones modified. As the writer suggested in a recent symposium on the use of projective tech-

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

niques for the assessment of normal personality development (5), several existing projective media appear particularly amenable to modification for theory-oriented personality studies. For example, specially-constructed sets of TAT-like pictures can be employed to test a variety of theoretical hypotheses. Blum's use of his Blacky Pictures to test selected psychoanalytic hypotheses is an example of this (1), as is the writer's employment of specially-constructed pictures to test Social Learning Theory (3, 4, 13). Incomplete sentences and stories, too, can easily be devised to elicit behavior relevant to specific theoretical hypotheses. Semistructured imaginative play situations can be used in a similar fashion. When existing projective media cannot be modified to fit research needs, new techniques will have to be devised. In theory-oriented research, however, the construction of these will be dictated by the problems under consideration rather than a desire to develop or use a new projective technique *per se* as has frequently happened in past research.

A second problem inherent in the use of projective techniques in personality development research is that of the categorization of responses. Since projective techniques are designed to elicit relatively "free," unstructured responses, these responses must be subsequently categorized by the researcher in some fashion. In past research, projective test responses have usually been analyzed according to a hodge-podge of empirically-derived and theoretically-unrelated categories. "Sign approach" methods for analyzing Rorschach protocols typify this approach. Such empirically-derived systems can seldom be used to categorize responses evoked by techniques other than those for which they were originally devised. As a result, data analyzed in such a manner usually lack generality. These data can neither be directly compared with data gathered by other personality assessment methods, nor can they be readily related to existing personality theories. Future personality research should include methodological studies concerned with the development of theory-oriented categorizations of responses applicable to data elicited by a variety of personality assessment methods. Ideally, such systems of categorization should be equally suitable for objective, subjective and projective responses when possible. Examples of existing methods of analysis which might be used in this fashion are Raimy's PNAv method developed within the framework of self-concept theory (12), Dollard and Mowrer's Discomfort-Relief Quotient influenced by Hullian learning theory (8), and the writer's Social Reinforcement Index developed within the context of Social Learning Theory (7, 13).

At the present time our knowledge of the areas and levels of behavior elicited by various methods of personality assessment is far from complete. Future personality development research should be facilitated to the degree that *cartographic studies* of existing assessment methods are a part of this research. These investigations could systematically map out areas, levels and frequencies of behaviors elicited from children of various ages by each assessment technique. Ultimately, this research should answer, among others, the following questions: (*a*) To what degree does a given technique cover the same behavioral areas with children of various ages? (*b*) At what level of expression is this behavior elicited? (*c*) Does the method produce an adequate sampling of this behavior to insure stability of measurement? (*d*) Is the range of this behavior great enough to allow sufficient intersubject and intrasubject variability? Such information should be helpful, not only to evaluate the usefulness of existing assessment techniques as research tools in personality development research, but also to determine areas where new methods must be devised. Cartographic information of this nature should ultimately open new avenues of research.

When suitable methods are available, for example, intraindividual covariational studies of children's objective, subjective and projective behavior will be possible. Such research will, of course, require assessment techniques covering comparable areas of behavior at these three levels of expression. When techniques are available, a variety of developmental studies are feasible. For example, while a number of recent investigations have dealt with the acquisition of social needs in children, there have been no studies, to the writer's knowledge, on changes in levels of expression of these needs as children mature. Such changes should be an integral part of a child's socialization. He comes to learn that the overt expression of certain needs is socially acceptable while a similar expression of other needs is not. Studies of adult subjects have indicated that correlations between objective and projective expressions of socially acceptable needs differ predictively from those of socially unacceptable ones. Future personality development research could study changes in such correlations in children's behavior with age, as well as factors producing these changes. Other investigations of a similar nature could be conducted in respect to the development of ego defenses in normal children.

The discussion so far has made no distinction between longitudinal and cross sectional investigations of personality development. Large-scale, long-term longitudinal projects such as those currently underway at several child research centers throughout the country are faced with one major problem not found in cross sectional research. This is due to the fact that, by the time large-scale longitudinal programs are organized and data are collected for several years, original concepts and methods of measurement are often obsolete. Researchers conducting longitudinal projects are continually faced with the decision of whether to introduce new concepts and methods in the study as it progresses, thereby sacrificing comparability of data collected from year to year, or to retain original methods and concepts when better ones are available. Obviously, there is no clear-cut solution to this problem.

One partial solution is what might be called *serial longitudinal research.* Long-term research projects following the development of children from birth to adulthood could be divided into a series of smaller studies, each covering consecutive two- or three-year periods. This design would allow continuity of subjects for the total investigation. In addition, some of the more durable concepts and methods could be retained throughout. When necessary, however, new methods and concepts could be introduced whenever a new short-term study was started. In this kind of program, each of the smaller studies could stand by itself as an independent investigation directed toward one or more problems germane to the period of personality development under consideration. In addition, data gathered from each of these studies could be used to make predictions concerning the subjects' behavior during later periods of personality development. Serial predictions of this kind, when checked against predicted-to behavior, should provide useful information on factors producing continuity or change in children's personality development as they mature.

Some of the studies might profitably focus on *transitional phases of social learning*. In our culture, while there are many periods of a child's life when his environment is relatively stable, there are also times when he is faced with rather drastic changes in his daily experiences. At these times, he encounters situations requiring the rapid learning of new social behaviors and the discarding of old responses when they are no longer appropriate. An example of one of these phases is the beginning of nursery school. The child is suddenly required to leave the stable, secure and familiar surroundings of home and mother for new, and oftimes threatening, experiences with strange children and nursery school teachers. This and other transitional periods should be particularly fruitful ones to investigate with intensive, short-term longitudinal studies; at these times the effects of specifiable environmental changes on social learning and personality development should be highlighted, affording the developmental researcher valuable opportunities for observing and measuring those relationships.

So far, this paper has discussed several methodological problems of personality development research. The remainder of the paper will be concerned with one basic theoretical problem, the generalization of social reinforcements. To the degree that personality development is a product of social learning experiences, the question of how social reinforcements generalize is an important one for personality development research. How do the effects of a given social experience determine a child's expectations and behaviors in subsequent situations? When a child experiences social rewards or punishments for his behavior in one situation, it is assumed that the effects of these reinforcements generalize to other situations in proportion to the degree of similarity between the original and subsequent situations. At the present time, however, we have little knowledge of factors determining psychological similarity in social situations. There is a paucity of information on the types of generalization operative in children's ongoing social behavior as well as the dimensions determining this generalization.

Several types of generalization may play a role in social learning and personality development. First, social reinforcement generalization may sometimes be a function of the *physical similarity of the stimuli* in the situation. Classical conditioning studies, e.g., those of Hovland (9, 10), suggest that this type of generalization is a reliable phenomenon in the laboratory. Whether it is important in children's social learning, however, must be determined in future personality development research.

A second kind of generalization which may influence children's social learning is that mediated through *similarities in conceptual-language structure*. Certain studies on semantic generalization and concept formation are relevant here. So too are some investigations on the development of prejudice. Few of these studies, however, have been primarily concerned with this type of generalization in social learning. In this respect, conceptual-language structure may be one important factor in personality development by functioning as a kind of perceptual-cognitive filter determining the effects of social reinforcements and their generalization. Until the problem is investigated directly, however, its role in social reinforcement generalization is an open question.

Social reinforcement generalization may also be a product of the *functional similarity of social reinforcing agents*. This has been suggested in psychoanalytic theory in such concepts as "father substitute" and the like. Unfortunately, most of these concepts are vague. They neither specify the particular characteristics of reinforcing agents producing reinforcement generalization, nor situational conditions conducive to its occurrence. Future personality development research can gain leverage on this problem in at least two ways. First, preliminary studies can be conducted determining the characteristics which children of various ages impute to specific reinforcing agents in their lives. Projective materials may be used in this research or actual persons may be employed as stimuli. The writer has found the latter useful with college students (6), but has not yet employed it with young children. Second, experimental studies can be devised in which the characteristics and functions of reinforcing agents are systematically varied, and resultant changes in children's behavior measured.

Social reinforcement generalization may also be mediated through the *similarity of an individual's social goals or needs*. Social behavior is characteristically goal-directed. A child, as the result of past social learning, develops certain needs and attempts to satisfy these through the attainment of corresponding social goals. When this goal-directed behavior is rewarded or punished, the child's expectations for future social reinforcements in respect to this behavior are altered. This experience may also alter the child's expectations and behaviors in respect to other similar social goals. The amount of this generalization would be expected to decrease with decreased similarity between the goal of the original behavior and other goals. Thus, for example, if a child is punished or rewarded while attempting to gain academic recognition, the effects of this reinforcement should generalize more to his behaviors directed toward other similar goals, such as social recognition or physical skills recognition, than to his behaviors concerned with dissimilar goals, e.g., the attainment of affection. Original research by the writer (4) and subsequent investigations by Jessor (11) and Chance (2) have demonstrated that such a process does occur in some social learning situations. To date, however, research on this phenomenon has been limited to adult subjects. Future child development research might include both studies dealing with this type of generalization in children's social behavior *per se* and investigations determining the role of this process in the original development of children's needs and social goals.

The four types of generalization which have been discussed do not, of course, exhaust the ways in which effects of social reinforcements may generalize in children's social learning experiences. Other types of generalization may also exist. Some of these may be identified in future animal research or classical learning and conditioning experiments. Still others may be elaborated in studies of complex social behavior. Regardless of how they are originally identified, however, it will be the responsibility of future personality development research to determine ultimately their respective roles in children's social learning and personality development.

REFERENCES

1. BLUM, G. S. A study of the psychoanalytic theory of psychosexual development. *Genet. Psychol. Monogr.*, 1949, 39, 3-99.
2. CHANCE, JUNE. Generalization of expectancies as a function of need-relatedness. Unpublished doctor's dissertation, Ohio State Univer., 1952.
3. CRANDALL, V. Induced frustration and punishment-reward expectancy in thematic apperception stories. *J. consult. Psychol.*, 1951, 15, 400-404.
4. CRANDALL, V. An investigation of the specificity of reinforcement of induced frustration. *J. soc. Psychol.*, 1955, 41, 311-318.
5. CRANDALL, V. Projective techniques in child development research. Paper read at Amer. Psychol. Ass., San Francisco, Sept., 1955.
6. CRANDALL, V., & BELLUGI, URSULA. Some relationships of inter-personal and intra-personal conceptualizations to personal-social adjustment. *J. Pers.*, 1954, 23, 224-232.
7. CRANDALL, V., & BELLUGI, URSULA. Rating personal adjustment through an analysis of social reinforcements. *J. consult. Psychol.*, 1956, 20, 49-52.

8. DOLLARD, J., & MOWRER, O. H. A method of measuring tension in written documents. *J. abnorm. soc. Psychol.*, 1947, 42, 3-32.

9. HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. genet. Psychol.*, 1937, 17, 125-148.

10. HOVLAND, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, 51, 279-291.

11. JESSOR, R. The generalization of expectancies. *J. abnorm. soc. Psychol.*, 1954, 49, 196-200.

12. RAIMY, V. C. Self-reference in counseling interviews. *J. consult. Psychol.*, 1948, 12, 153-163.

13. ROTTER, J. B. *Social learning and clinical psychology.* New York: Prentice-Hall, 1954.

A PROGRAM OF RESEARCH ON THE DIMENSIONS AND ANTECEDENTS OF EMOTIONAL DEPENDENCE

Jacob L. Gewirtz[1]
The University of Chicago

The problem set for this paper is to outline a program of research which focuses upon an assumed continuity between early childhood conditions and those consequent social behaviors of children which typically have been labeled "emotional dependence." In particular, emphasis is placed upon behaviors which are employed apparently to gain such positive social responses from other persons as those involving attention, approval, affection, reassurance, and nearness, and upon the early conditions under which they develop. While it is probable that knowledge of the processes underlying emotional dependence will be critical to the understanding of a substantial portion of the social behavior of children and adults, very little is known either about the dimensionality or the antecedents of that behavior class.

The principal focus of this paper is upon child behavior; however, the major problems raised may be considered a subclass of general behavior problems and hence continuous with questions characteristically raised about adaptive behaviors exhibited in simple settings both by infrahumans and humans. Since the theories of social development which do not have their origins in such basic considerations of learning appear to approach the earliest conditions for the development of dependence in a less articulate manner, selected descriptive concepts and postulates which stem from work on general behavior problems are emphasized in this paper. Relatively little emphasis is placed upon concepts at other than simple levels of analysis (e.g., personality characteristics of the mother, her acceptance of the child, social class membership), which appear to be employed in many theories and researches on children's dependence, and index conditions which would qualify early dependence learning. However, it is expected that the reader will bring to mind such concepts, theories and research findings as they are required.

A systematic ordering scheme would appear requisite to any research program of intended generality and significance. It is preferable that such

---

[1] The writer is grateful to Mr. D. M. Baer for his valuable aid in the preparation of this paper.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

a scheme be susceptible of simple, operational statement and include a neutral, parsimonious set of terms, yet be capable of flexible application to a variety of data, procedures and samples of subjects (*S*s). Such a system could generate hypotheses, as well as articulate with other systems of language and theory. These are rather general criteria. Additional special criteria are suggested by the problem at hand: In particular, it would seem fruitful to be able to analyze the apparent complexities of emotional dependence into rather skeletal terms so that they may be related in a simple manner to the fundamental mechanisms operative in the developing child. Thus the ordering system required must at once be capable of reducing to simple terms the complexities of the behaviors which characterize emotional dependence and of positing basic acquisition processes which may be applied to order the multiplicity of environmental conditions to which children are typically subject. And if, in addition, the properties of this ordering system were quite readily reconciled with extant concepts and theories, considerable advantage would accrue.

The basic definitions and principles in the sections which follow are intended to meet the criteria enumerated above, and hence prove useful in researches of the type required; and, further, they can serve to develop some of the background issues, both of terminology and substance, to provide the frame of reference in terms of which the special problems involved in such a program of research can be considered. It is hoped that the reader will bear with us in the consideration of the familiar concepts which will be presented in the sections which follow. The concepts and assumptions employed are drawn heavily from those of the descriptive behaviorists (6, 8). The basic paradigm employed, it will be seen, is that of *instrumental* or *operant conditioning*.

BASIC DESCRIPTIVE CONCEPTS

The Operant

An instrumental or operant *response* is any "voluntary" (non-reflex) response which is originally emitted more or less independently of identifiable stimuli and which can manipulate the environment to produce changes which are of consequence for the organism. While this definition embraces much of human behavior, in this paper interest will center upon such responses as those employed for or maintained by the attention, approval, affection, proximity, caresses, and similar characteristics of another social being. While a particular response may be "spontaneous" in its initial occurrence (e.g., the gross movements of the infant), as a consequence of reinforcement it very soon can become associated with particular stimuli, such that it becomes more likely to occur in their presence. A particular stimulus which "sets the occasion" for a particular response has been termed a discriminative *stimulus* (S^D) or cue, and the response involved has been termed a disciminated operant. It is assumed that discriminative stimuli constitute the major portion of the controlling environment for an organism. With reference to responses characterizing emotional dependence, discriminative stimuli are typically afforded by people, and thus may be termed *social*.

The Reinforcer

Any stimulus event which follows a response and affects certain of its aspects, e.g., its rate of emission, amplitude, or latency, is termed a *reinforcing stimulus* or *reinforcer*. Reinforcing stimuli are of two basic types: *positive* reinforcers strengthen responses when presented; and *negative* reinforcers strengthen responses when removed. Many of the stimulus events which afford such control over responses are supplied by people (e.g., affection), and may be termed *social* reinforcers. The concepts of response, discriminative stimulus, and reinforcer appear to allow the ordering of the phenomena covered by the rubric emotional dependence. Social behavior in general has the appearance of being directed; that is, it may be analyzed as employed to attain some happening or stimulus "commodity" from other people, like attention or approval. The concept social reinforcer directs attention to this characteristic, insofar as it focuses upon the stimulus events which yield control of responses when their presentation is made contingent upon them.

Drive

However, it is frequently observed that even in the presence of discriminative stimuli setting the occasion upon which a response shall be reinforced (by the contingent supply or removal of a reinforcer), the response is not emitted, or, if it is, the contingency of the reinforcer upon the response yields no control over it. This is tantamount to stating that on such an occasion the reinforcer is not operative as such, i.e., that there exists variance in its efficacy. This variance is generally attributed to a condition of *drive* holding for the organism at the time in question, particularly when it is demonstrated that the stimulus event in question acquires reinforcing power for the organism only subsequent to his recent deprivation of it, and loses this power subsequent to his satiation by it. In general, such stimulus events appear to possess this characteristic innately, e.g., food and water. Other stimulus events which possibly meet the same criterion may be some whose reinforcing value was acquired through a history of conditioning.

At the behavioral level, the criterion for what is and what is not a drive operation is arbitrary, i.e., one of definition and convention. In general, treatment variables are classified primarily in terms of their effects in behavior. Hence, for example, the drive label may be reserved for all nonconditioning operations which affect the efficacy of reinforcers, or it may be restricted to but a few operations which appear to have similar effects and, perhaps also, which are judged similar according to other criteria. So also, conditioning operations in general may be classified in a number of ways. The particular classification system one employs would depend on the behavioral effects he has focused upon, the data accruing from earlier work in the problem area, labeling conventions in that area, and the conceptual system (if any) within which framework he has been working. For much work, labeling operations and advancing hypotheses are indistinguishable processes.

While the concept of drive has a variegated and somewhat confused tradition, in this paper it will be taken generally to represent almost *any* operation, apart from a standard conditioning procedure,[2] which can be shown to affect changes in response output rates. Hence, where reinforcing stimuli are identified, a drive operation is taken to be one which affects the relative efficacy of a number of reinforcers. Since for the purposes of this paper it is convenient to emphasize reinforcers rather than responses, the latter working definition of a drive operation will be employed. Where reinforcing stimuli appear to show no variance in their effectiveness, the concept of drive is unnecessary, although the identity of such reinforcers might be of significance for a variety of problems. Emotional stimuli which produce "states" like those termed frustration and conflict also may affect the relative efficacy of reinforcers. Hence the conditions producing such stimuli may be termed drive operations as well, for present purposes.

Conditioned Reinforcers

A stimulus event which is innately reinforcing is termed a *primary reinforcer* (S^R), and if its deprivation affects its reinforcing properties, then such a deprivation is ordinarily taken to define a state of *primary drive.* A discriminative stimulus which acquires a reinforcing character only subsequent to a history of pairing with a primary reinforcer, typically (but not necessarily) with the organism under a state of deprivation for that reinforcer, is labeled a *conditioned or secondary reinforcer* (S^r). It is usually assumed that the effectiveness of a conditioned reinforcer varies with the deprivation appropriate to the primary reinforcer to which is has been conditioned.

There exists the possibility of pairing a conditioned reinforcer with more than one primary reinforcer. This stimulus event has been labeled a *generalized reinforcer* by Skinner (8), and is assumed to be effective subsequently under any one of the deprivation conditions appropriate to the primary reinforcers to which it has been conditioned. Since at least one appropriate state of deprivation is likely to prevail for an organism at any given time, the generalized reinforcer would be effective at almost any time. Further, it appears reasonable to expect that typical child-rearing practices may condition certain discriminative stimuli to a certain number of primary reinforcers under their appropriate deprivation states; hence this construct may have much relevance to emotional dependence learning.

---

[2] The criterion used here is that such an operation may be instituted without requiring more than at most a very few further trials to produce its effect. Such operations typically can be instituted at any moment for an organism, although the operation itself may not be independent of a history of conditioning (e.g., possibly the deprivation of a conditioned reinforcer, or breaking into an existing chain of responses in a manner which prevents completion of the chain, as in frustration).

Conditioned Drives

In the past, the apparent potency of a number of stimulus events which may meet this description of generalized reinforcer has supported the usage of terms like *social, secondary* or *conditioned drives,* e.g., for attention, approval, or the company of people. Under the present definitional scheme this usage would be in error if the effectiveness of these reinforcers could be shown to be under the control of all pertinent primary deprivations; then the phenomena properly would be referred to the characteristics of generalized reinforcers. However, it appears reasonable to propose that if the deprivation of a conditioned or a generalized reinforcer would increase its own effectiveness as a reinforcer, and if its satiation would reverse that effect, then such a deprivation may be defined as a *conditioned drive* operation. It is important that *this* definition turn essentially upon this notion of deprivation-satiation operations for the conditioned reinforcer in question; for it is usually observed that the effectiveness of a conditioned reinforcer can be referred to the organism's deprivation of the particular primary reinforcer involved in the conditioning.

If it can be demonstrated that this criterion for conditioned drive is met when children are deprived of particular social reinforcers, then a construct of great potential power for ordering social behavior would result. Since under typical social conditions many social reinforcers are not available to an individual for moderate or long periods of time, the apparent prevailing potency of these reinforcers could be assigned to their empirically accessible deprivation properties. Further, the fact that the deprivation of some social reinforcers might produce drive effects while the deprivation of others might not, would imply that certain qualitative differences in socialization procedures probably maintain. The membership of the events in each of these two classes of social stimuli would provide significant clues to these differences.

The convention about drive maintained in this paper does not require that the responsiveness of the potency of a conditioned reinforcer to its own deprivation be the only criterion defining conditioned drive. As stated earlier, drive is invoked here as a construct only to account for variance in the efficacy of a reinforcer. There may be a variety of operations distinct from standard conditioning operations which affect the potency of reinforcers in a similar manner, and drive as a construct is intended to include these, for present purposes. Insofar as any of these operations reflect the play of some variable established as such through a specifiable history of conditioning, it appears appropriate to refer to the application of that variable as a conditioned drive operation, if only to maintain a certain consonance with traditional usages of the term drive.

This concludes a brief outline of the conceptual scheme, and the hypotheses it generates, which is to be applied to the problem of the dimensionality and the development of emotional dependence in childhood. Finally, it should be noted that this system is intended to be open to modification and addition; where new constructs prove necessary to account for variances not accounted for by the basic constructs employed, in principle they *must* be developed.

PERTINENT THEORETICAL STATEMENTS

The development of emotional dependence has been the subject of relatively little formal theorizing. Speculation has sometimes taken the form of extending to some aspects of dependence existing theoretical schemes which had proven useful in ordering relatively simpler learning phenomena (e.g., 1, 8), as is the case in this paper, and sometimes the form of describing and interpreting clinical case histories in terms of emotional dependence. While considerations stemming from many sources have been taken into account in the preparation of this paper, limitations of space do not permit their detailed examination.

THE BASIC PLAN OF RESEARCH

The overall research program is to proceed essentially from the simple learning and performance concepts set forth in earlier sections toward tracing the development of emotional dependence in children. The most fruitful application of these concepts demands that certain priorities be established among the many classes of research which they can generate. Initially, a two-pronged approach appears suitable: the *dimensional* program can begin with an examination of the characteristics of behaviors typically labeled emotional dependence; while the *developmental* program can focus upon their antecedents, both at simple and complex levels of analysis.

The Dimensional Program

A possible strategy in the research on early dependence learning would be first to establish a catalogue of the characteristics of emotional dependence, the details of which might then be employed in the researches of the developmental program to focus effectively upon the events which typically operate in that early learning. This program would begin with an enumeration of the potentially limited number of social reinforcers, of responses employed for those reinforcers, and of object persons who can provide the reinforcers, all of which characterize the process in young children. Next studied would be the relationships holding among responses, reinforcers, and objects, and the distribution of hierarchies of efficacy of the reinforcers for various populations. Given the beginnings of this catalogue, this program finally would add developmental level, defined in terms of age, as one of the basic parametric variables with which it will deal, thereby ordering emotional dependence as a developmental process. While reflecting the cumulative history of the organism, age provides a crude index of a number of developmental processes, and represents a dimension correlated to stages in development at which qualitative changes occur. The mechanics of these variables are dealt with in the program which is described in a later section.

Some preliminary work relevant to the dimensional problem can illustrate some of the above points. In an investigation of some parametric and dimensional aspects of the attention-seeking process, it was found possible to distinguish nine responses which were assumed to be employed for an adult's attention, and some perhaps also for his approval (2). These could be reliably scored, and entered into a pattern of relationship with the independent variables. Further, it was found that the six responses which were verbal in character clearly defined a factor in a factor analysis (3); and that when the distribution of frequency scores for each response was dichotomized about zero frequency, and the two responses with the most zero scores were combined, a five-item Guttman scale was produced (4). The responses composing the scale, according to their increasing scale values, are: Comments (casually emitted), Questions (casually asked), Attention-seeking (urgent responses for overt notice from the adult), Help-seeking, and Permission- *or* Praise-seeking. Thus, it appears that a dimension of behaviors apparently employed for the reinforcer of attention, and perhaps also for approval, has been identified, and can serve as a first approximation for ordering dependence responses. In addition, age yielded almost no reliable relationships with the nine responses in the four- to six-year age range employed. However, when Heathers (5) compared the free play behaviors of 20 two-year-olds and 20 four- and five-year-olds, he found that clinging and affection-seeking decline with age relative to attention- and approval-seeking, and that emotional dependence on adults (nursery teachers) declines with age relative to emotional dependence on other children. The writer has found that children age four to six exhibit relatively greater frequencies of responses for attention to adults of the opposite sex as themselves than to adults of their own sex, with the difference much more pronounced in boys than in girls (2).

These results indicate that the goal of isolating and cataloguing a variety of responses, reinforcers, and objects involved in dependence behaviors, and of stating this catalogue as a function of developmental level, is reasonable.

Some further experiments were conducted under the writer's auspices to evaluate the effects of reinforcers which, in the setting of the experiment, presumably appealed to the reinforcer of approval. They were carried out also to provide, in the differential susceptibility of *Ss* to such reinforcers, a dependent variable which might yield considerable leverage upon the characteristics of emotional dependence. The reinforcers were found to exhibit a surprising complexity: for while it was established that a number of simple responses in children of various ages (e.g., bar-pressing, marble-rolling, letter-naming, word-emission) could be controlled effectively by the casual use by the experimenter of terms like "good" and "uh huh" (which control was in the main unrecognized by the *Ss*), it was also found that apparentaly identical control (also unrecognized) was achieved by using for a reinforcer a sound (that of a penny dropping into a bottle). Unlabeled learning or learning without awareness may provide a prototype for early identification learning in children.

The Developmental Program

The developmental program describes researches designed to supplement the dimensional program by establishing the laws underlying the development and subsequent efficacy of the generalized reinforcers, and possibly the conditioned drives, assumed to be involved in emotional dependence. In effect, this requires the eduction of some specific laws of learning to the special case of the developing human, and the initial focus upon the earliest dependence learning susceptible to experimental investigation. In order to accomplish this, a two-pronged developmental research program designed to converge more or less simultaneously upon different levels of problems is outlined: The first will investigate the parameters of generalized reinforcers, employing relatively simple (i.e., stripped down) learning situations to elucidate the processes involved; and the second will deal (perhaps at several levels of analysis simultaneously) with the actual details of the presumably complex learning of emotional dependence in the young child. While the latter is emphasized, clearly the two subprograms are interdependent.

The Basic Characteristics of Conditioned and Generalized Reinforcers. Researches of the first subprogram would identify the variables which determine the strength of conditioned and generalized reinforcers, and then study those variables as parameters. Thus, there would be studied the effects of such training variables as: number of pairings of conditioned with primary reinforcers; schedule of such pairings; temporal relation between conditioned and primary reinforcer; magnitude of primary reinforcer; and number and type of primary reinforcers, and of deprivation states involved; and different effects on reinforcers conditioned by positive and negative primary reinforcers. In addition, selected performance variables (e.g., different deprivation conditions) under which conditioned reinforcers are operative, as well as the effects of the deprivation of conditioned reinforcers, would be studied. The basic information provided by such researches could serve as one class of hypotheses for the study of the details of the child's emotional dependence learning. Fortunately, research on basic learning continues to produce studies of this class.

Generalized Reinforcers and Emotional Dependence. Departing from the simplest level of analysis that the first subprogram represents, one can move as freely among levels of complexity as is required by the problem at hand. Stimulus events more social than, e.g., the lights or buzzers employed in studies of the first subprogram may be investigated as discriminative stimuli and conditioned reinforcers (e.g., physical appearance, attention, or approval of a caretaker), as may be relatively social responses (e.g., "Look at me!"). In addition, the variance in the potency of a number of presumed generalized reinforcers already effective for children in the early years, can be studied in relation to specific primary deprivations, some of which were involved presumably in early dependence learning. Other deprivations can be of the generalized reinforcers themselves, to determine whether the effects produced are similar to those produced by primary reinforcer deprivations, or by other drive operations. If such operations were to meet the criteria, they would be termed conditioned drives. The results of such researches could be of considerable utility, for they might point to at least two qualitatively different types of early learning, one producing generalized reinforcers alone, and the other producing conditioned drives as well. If this possibility eventuated, the identities of the events falling into these two classes, as well as the identities of the primary deprivations to which they might be sensitive, would yield valuable clues about early childhood learning.

At times, loosely-defined dimensions such as "rejecting-accepting," "warm and nurturant—cold and neglectful" and the like have served as antecedents of a variety of social behaviors in theories and descriptive laws; they have described long term patterns of behavior of caretakers towards their children. It could prove useful to the program outlined, and perhaps serve to articulate it with a number of extant theories and researches, if some of these gross concepts were examined, and the reinforcement histories involved (e.g., schedule of reinforcement, identity of reinforcers) postulated.

As a unit of the dimensional program, an attempt to determine the distribution of hierarchies of reinforcer efficacy is to be made, and groups of *S*s found homogeneous in this respect established. Given such information, it would be possible to attempt to establish the identities of such groups according to criteria as, for example, might be suggested by the conceptual scheme outlined, or by others. Communalities in antecedents could be identified in this manner, and information valuable for the developmental research program obtained.

Concurrently, researches on an infrahuman level may be conducted to examine the effects of postulated variables for the special case of human infancy. In particular, research on this level could be useful in studying variables too extreme or too vaguely understood to be well implemented with humans, as well as for certain variables involved in long term longitudinal studies.

A Possible Experimental Program with Infants. Finally, a variety of researches on human infants under such caretaking conditions as would be provided by a highly controlled institutional or home setting would become feasible and desirable as data accrued from the dimensional and developmental researches. These would be either in the form of experiments which would attempt to produce particular effects subsequently in the child's pattern of emotional dependence (reinforcer and response hierarchies) by varying selected aspects of the caretaking process; or they might employ systematic observational techniques to assess the learning contingencies for the child, which could be referred subsequently to his pattern of emotional dependence.

The experimentation would be aimed at producing a prototypic "model dependence" in young children. Both positive and negative reinforcing aspects of caretaking may be employed as variables. Selected stimulus events involved in or attached to the caretaking person may be set into a variety of contingencies with these aspects of the caretaking process; and selected aspects of this process may be made contingent upon different responses emitted by the child.

By postulate, such procedures will imply consequences in the later behavior of the children to whom they are applied, and the experimenter may find that he has control of the laws governing some major dimensions of social interaction. Researches initially would be restricted to the use of contingencies holding between socially trivial stimuli such as lights or buzzers, and socially trivial responses such as head rolling, rocking, or playing with a toy. The range of control of the caretaking process must be established cautiously. In instances where it is difficult to establish clearly the nature of the variables to be applied, or the most efficient method of applying them, the problem may be referred to analogous experiments on animals. Responses such as head rolling or rocking, aimed at securing "model" reinforcers such as lights or buzzers, either from people of from a simple mechanical apparatus capable of their delivery, would be studied. Hence it may be possible to speak of "light-seeking" as a "social" behavior or simply as a general behavior. The effects of other parameters, such as the scheduling of the contingencies, the differential aspects of stimuli, caretaking, and caretakers involved in those contingencies, and such general parameters as age, or pre- and postverbal learning, may be applied both for the general case and for the "social" case.

These relationships need not be restricted to the "model" level described above, for neither the stimuli nor the responses involved need be trivial. Such events as the attention, approval, or affection of the caretaker may be the principal stimuli involved, and responses of social significance, viz., attention-, approval-, and affection-seeking, may result. Validation of the model described must be affected cautiously, and it may be possible to conduct such researches only in a late stage of the program.

SOME GENERAL PROBLEMS

It is expected that many of the behaviors of young children may vary with developmental level. Any research program aimed at developmental functions would do well to anticipate such of these as appear relevant. For example:

1. The hierarchy of reinforcers for the young child may vary with developmental level. A portion of this variance may be attributed finally to variables investigated in the programs described here, while another portion of the variance may result from very general variables at present not of immediate relevance to those programs, e.g., the types of discipline employed by the parent at various ages. Such effects may be reflected in the proposed catalogue of reinforcers, and could be evaluated there as such.

2. The efficacy of a given set of learning procedures may vary with developmental level. For example, it appears likely that preverbal and postverbal learning differ in several qualitative aspects. Thus, it is possible that two individuals to whom apparently identical conditioning procedures have been applied will show different sensitivities to the reinforcers involved, simply because the procedures were instituted at different developmental levels.

Other problems or issues of a general nature which appear relevant to the program are:

3. It is possible that certain inevitabilities of human caretaking will produce some response-reinforcer correlations of a very general occurrence. For example, it may be discovered that attention is an inevitable reinforcer for the response of talking. Hence, it may be difficult to account for such correlations initially, since there may be relatively little variance apparent in the population, and experimental control may be impractical. Consequently, the catalogue of reinforcers may appear to contain "unexplicable" relationships.

4. The acquisition mechanics utilized in conceptualizing this problem allows a wide range of possible stimulus events to meet the criteria defining the development of generalized reinforcers pertinent to dependence behaviors. However, certain aspects of the typical child-rearing situation yield a number of expectations about the identities and interrelationships of these stimulus events, which may operate to reduce substantially the number and range of the significant generalized reinforcers typically conditioned. Hence, these expectations must be examined and evaluated for their implications.

For example, it is thought that before a child is reinforced by his caretaker, certain stimulus characteristics of that caretaker typically will be made apparent. In the general case, it may be anticipated that caretaking is *most* probable from a caretaker who is at once (*a*) close at hand, (*b*) attentive to the child's state of need, (*c*) approving of his request for care, and (*d*) affectionate in the delivery of that care; on the other hand, caretaking is least probable from a distant, inattentive, disapproving, disliking person. Hence, there can be postulated a continuum of the probability of caretaking as a function of the differential presence of these stimuli, which (according to the acquisition mechanics utilized) implies consequent sets of hierarchies of these same stimulus events as generalized reinforcers controlling dependence responses. The following appear to be some major hierarchies which may result from typical caretaking conditions:

i. The proximity and attention of the caretaker may be the stimuli most strongly correlated with the caretaking process. This is probable particularly in the case of the very busy caretaker, who can be near the child only to see to his needs, and must be elsewhere at other times. However, in an environment of very limited space, the proximity of the caretaker may be inevitable and continual. If, in this situation, the caretaker is very busy, then her attention may be the single stimulus most strongly correlated with caretaking.

ii. The proximity and attention of the caretaker will not be so strongly correlated with caretaking in a situation where she has greater leisure, for then her presence and attention will be available on many occasions which need not be caretaking occasions. Thus, it is likely that there will be times when the child may receive care of a minor nature. In such situations, it is anticipated that other characteristics of the caretaker, e.g., her approval or affection, may fulfill the criteria of discriminative stimuli more adequately than will her proximity or attention.

THE METHODS OF RESEARCH

The research program that has been proposed is both diverse and intensive. Hence, ideally it would require highly detailed investigations of a considerable number of problems. It is not appropriate here to detail experimental designs for these researches: to some extent they are a function of the data accruing early in the program, and hence must await those data. Consequently, it is practicable to outline only the basic methods which will probably be required in the overall program.

It is recalled that an initial problem is to recognize and identify the social reinforcers which characterize emotional dependence behavior. The definition of a reinforcer provides an immediate criterion for this problem: it is a stimulus event which, if made contingent upon some response, yields control of the increasing strength of that response. However, this definition should be entertained only simultaneously with the general expectation, noted earlier, that some reinforcers may be pertinent more to certain responses than to others. In general, this initial problem of defining the responses and reinforcers comprising the class of emotional dependence behaviors requires some type of standard situation, within which it is possible for *Ss* to emit responses upon which reinforcers may be made contingent. This approach, with suitable variation, allows for the definition of social reinforcers, and for some assessment of their possible correlations and hierarchies.

The Baseline Method. The introduction of this method into psychology is due primarily to the work of Crozier and Skinner. Its logic is simple, and may be applied efficiently to experiments required in the program outlined. In terms of statistical design, the method essentially represents a treatment by subjects design (perhaps including a variable representing successive units of time or trials), in which a highly dependable response variable is employed.

In principle, the logic of the method is as follows: if the rate of output of any instrumental response can be elicited dependably under standard conditions, those data would provide a baseline against which the effects of treatments could be evaluated efficiently. When a stable rate of output is attained, for the purpose of an experiment a major portion of output variance, perhaps with the exception of variance due to differences between *Ss*, is under experimental control. Thus, a highly efficient experimental method is provided, one which generally requires small numbers of *Ss*. Considerable exploratory work typically is required to establish conditions under which the highly reliable dependent variable is generated.

The nature of the response selected would depend on the given problem. A simple response may be employed, even in a program whose focus is the study of complex human processes. The experimenter's (*E*'s) logic might content him with the laws involving the simple response, or he might choose to validate some of those laws against laws which involve the more complex behavior; and then, if a high degree of correspondence were found to obtain, he could continue to employ the simple response as his most efficient vehicle for gaining leverage on the complex behavior class. Alternatively, he might employ the complex response from the beginning.

The steps in the typical experiment are very much as follows: First, a response (e.g., bar-pressing, naming-things, attention-seeking, etc.) is selected. Then, a dependable rate of output is established under a given set of conditions, such as that generated by a particular schedule of reinforcement, or set of instructions. If the dependent variable so generated can take a wide range of values, it can provide the required baseline. After the constant rate of response output has been developed in *S*, a treatment is introduced, and the response rate observed. When the rate has become stable under the treatment, the reversibility of the process can be examined, as can the efforts of different degrees of the experimental variable, allowing perhaps for the examination of the complete range of the variable in one *S*.

The number of repetitions necessary to establish its reliability would depend on the magnitude of the effect and the consistency of the change in rate of response from treatment to treatment. The experiment is repeated on a number of *S*s sufficient to establish that the effects of the treatment are significantly large with respect to individual variability. Few *S*s are required if the output records indicate that the response effects vary in the same way from *S* to *S* (and differ perhaps only in terms of a constant effect, e.g., a *S*'s level of responding).

Some Illustrative Techniques. To date, employing the baseline logic, there has been developed to a nearly satisfactory stage a method for defining social reinforcers, when behaviors to attain them are channeled into one of two motor response outlets. The child (*S*) is introduced to the setting as to a game, the central aspect of which is a toy displaying two holes down which marbles may be dropped (to return at the base of the toy). *S* is given a supply of marbles, shown briefly how to play the game, and allowed to establish baselines for the use of each hole. Subsequently, *E* chooses the nonpreferred hole, and makes some social reinforcer contingent upon its use. This contingency may be announced to *S* or not, depending upon the formal characteristics of the reinforcer utilized. Consequent shift from baseline use of each hole indicates the degree of control that the reinforcer involved can exert, for a given *S* on a given occasion. In principle, it is possible to articulate a variety of closely discriminated social reinforcers into this situation by tailoring the set appropriately.

For example, *S* may be introduced to the toy, and to the recording apparatus housed in an adjoining room. He may be told that the toy runs the recorder. After his baselines have been established, *E* may announce that he can stay and watch *S* play only so long as marbles are rolled down a certain hole; otherwise, he must see to tasks (e.g., tend the recorder) which requires his attention in the next room. Degree of use of that hole by *S* in excess of his baseline use of it assesses the potency of the social reinforcer involved in this set, which tentatively may be labeled as *attention involving proximity*. An alternative set, which appears to appeal primarily to *approval,* is constituted by *E* remaining in the room with *S* and, subsequent to the establishment of the baselines, pointing to one of the holes and remarking, "You know, this a the good hole." This set can be elaborated by *E* expressing approval subsequently whenever the hole is used (e.g., by saying "Good!," "Fine!," "Uh huh," or "That one was good"). An alternative set, appealing in its face characteristics to *achievement* as well as approval, can be constructed from the approval set by *E*'s counting aloud the number of "good" responses which *S* makes. Alternatively, achievement may be suggested more clearly by telling *S* that in this game he can get "points" if he wants, and then dispensing tokens, or making a tally mark whenever *S* has employed the correct response a sufficient number of times. Here the game element can be preserved by letting *S* discover the identity of the hole which yields points. Again, the ability to *internalize* approval or achievement may be assessed by E leaving the room after the approval set has been established, thus leaving S to maintain the correct response (if he will) alone. Processes such as *identification* may be assessed by establishing a set in which E, after noting S's baseline, himself takes a turn at playing the game (as a friendly demonstration), but restricts himself almost completely to the use of one hole. Consequent copying behavior on the part of S may qualify under a given experimental logic for the label of identification.

From these examples, it is clear that a large number of simple, natural sets can be constructed which have a face appeal to the variety of stimulus components typically thought to constitute the reinforcers involved in emotional dependence. Experience suggests that each S may be used again and again, under the same or different sets, if these occasions are separated in time by a few days. Hence in principle it is possible to establish efficiently and in fine detail many of the intercorrelations, equivalencies, and hierarchies of these social reinforcers. Since it has been found, to date, that children as young as four perform satisfactorily in such settings, these interrelationships may themselves be stated as function of a developmental parameter indexed by age. Such findings would have considerable relevance to a number of theoretical postulates about the development of these behaviors; for example, need a child be strongly responsive to approval before he will display behaviors labeled as identification, or achievement?

By the same token, a wide class of nonhistorical antecedent variables may be articulated into these experimental sessions, thereby establishing laws which may be basic to everyday social interactions. This possibility is quite reasonable when it is recalled that, while the experimental situation involved here is a skeletal, "stripped down" setting, nevertheless the stimulus events delivered as social reinforcers represent a fair sampling of the reinforcers utilized in life settings; similarly, the antecedent variables applied also may be drawn from a class of typical life events. For example, an S may be run under any set (i.e., for any social reinforcer) immediately after a period of social isolation. This variable can be affected in practical terms by isolating S in a room, with a toy to play with or a task to solve, and leaving him there for as long as is practicable. Any effects which this procedure will reflect in the typical potency of the reinforcer subsequently dispensed in the experimental session can be of great relevance in accounting for variance in behaviors for such a reinforcer in everyday life: whether a period of social isolation is conceived of as a condition of deprivation or as a change in the schedule of receipt of social reinforcers, still it parallels the fluctuations in the availablity of social reinforcers typical to an everyday social environment.

In a similar manner, many classes of social variables could be abstracted from complex environmental contexts and applied to these experimental sets, and their effects assessed in the consequent responsiveness of responses for the social reinforcers dispensed there. As examples, consider sex of the reinforcer-supplying adult relative to the sex of the child, selected primary deprivations, and the definable status groups (peers, etc.) from which spectators are drawn.

However, it must be remembered that in a less structured situation where the child can choose his "reinforcer-seeking" responses, it may well be expected that some responses may be more pertinent to some reinforcers than to others. To assess these possible reinforcer-response correlations, a standard situation like that employed in Reference 2, leaving the child free to show a wide range of responses, is required. (The above situation which channels all responses into a single motor response clearly is not applicable to this problem.) It is also possible to institute in such a situation, e.g., two or more levels of the availablity or supply of such reinforcers, and hence the interrelationships of the dependence responses emitted can be reflected against such a basic variable. Such comparison allows deductions about the essential nature of many of the reinforcers, e.g., their equivalencies or hierarchies under conditions of relative commitment of the adult as a source of reinforcers.

Other Procedures

The results of factor analyses may serve to demonstrate independent factors operative in the phenomena of emotional dependence, which possibly can be referred to distinct prior processes. Similarly, the use of some recently developed scaling methods for qualitative data, and in particular Lazarsfeld's method of latent structure analysis (7), may serve to define groups of *S*s which are homogeneous with respect to different hierarchies of generalized reinforcers; and if it were possible to establish the identity of these groups in terms of such criteria as distinct prior learning contingencies, then a number of new research avenues would be opened. The range of independent criteria on which such groups might differ is broad, and hence is best left to the reader's imagination.

SUMMARY

A research program is outlined for investigating the laws characterizing emotional dependence in childhood. In particular, emphasis is placed upon the identification of the dimensions operating in behaviors which appear to be employed to gain certain social responses from other persons (like those apparently employed for attention, approval, affection, reassurance, and the like); and the conditions under which they develop. The program proceeds both from simple learning and performance concepts and from such theories as are available for tracing the early development of emotional dependence.

Initially, a two-pronged research program is proposed: the *dimensional* program would catalogue instrumental behaviors, reinforcers and object persons which characterize the dependence process, and such interrelationships as may be discovered among them, stated as a function of age; and the *developmental* program would employ the catalogue to focus upon the etiologies of different emotional dependence patterns. In particular, these researches propose to focus systematically upon selected aspects of the human caretaking situation.

REFERENCES

1. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.
2. GEWIRTZ, J. L. Three determinants of attention-seeking in young children. *Monogr. Soc. Res. Child Develpm.,* 1954, 19, No. 2.
3. GEWIRTZ, J. L. A factor analysis of some attention-seeking behaviors of young children. *Child Develpm.,* 1956, 27, 17-36.
4. GEWIRTZ, J. L. Steps toward a scale for children's attention- and approval-seeking. In preparation.
5. HEATHERS, G. Emotional dependence and independence in nursery school play. *J. genet. Psychol.,* 1955, 87, 37-57.
6. KELLER, F. S., & SCHOENFELD, W. N. *Principles of psychology.* New York: Appleton-Century-Crofts, 1950.
7. LAZARSFELD, P. F. Latent structure analysis. In S. A. Stouffer *et al., Measurement and prediction.* Princeton, N.J.: Princeton Univer. Press, 1950. Pp. 362-472.
8. SKINNER, B. F. *Science and human behavior.* New York: Macmillan, 1953.

SOME RESEARCH PROBLEMS FOR DEVELOPMENTAL PSYCHOLOGY

EUGENE S. GOLLIN
Institute of Child Welfare
University of Minnesota

The function of this paper, as outlined by the organizers of the conference, is to present matters of high research priority in the area of developmental psychology. Such presentation is bound to reflect theoretical, methodological, and personal biases; it follows that many matters of high priority are thus excluded. It is with this recognition that the following considerations are submitted.

Reviewers of the field of child psychology have been relatively pessimistic of late with regard to its general status. General pessimism was expressed by Barker in the second volume of the *Annual Review*. He stated that the field was lacking in vigor (2). By the time the fifth *Annual Review* was prepared the problem was no longer one of the vigor of the field but rather was concerned with the lack of rigorous methodology and systematic theory (17). Other criticisms have been directed toward the apparent lack of integration between child psychology and general psychology, and the use of the child as a special subject for study in his own right (23).

The former concern operates both ways. For example, while some feel that child psychology has ignored general psychology, others assert that general psychology has been slow to take advantage of developmental principles (44).

The utilization of the child as a subject for study in his own right has applied value, but researchers involved in the pursuit of general principles have the responsibility of viewing their developmental samples within the encompassing framework of phylogenesis and ontogenesis. It ought to be

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

recognized that child psychologists are not the only psychologists who from time to time abandon this perspective. See for example Harlow's caustic comments on "rodentological learning theory" (10), Beach's remarks about the narrowness of species sampling in general experimental psychology (3), and Verplanck's call to observe more species and observe them under a wider variety of behavioral conditions (including ontogenetic periods) than those generally utilized in the laboratory (40). These criticisms are directed at general and experimental psychology and would indicate that much of the pessimism expressed about child psychology should be generalized to those fields.

The outline of a high priority research program is inherent in the dissatisfactions cited above. The theme of such research must be *change* in the sense that Sears has employed the term. Sears discusses two aspects of behavioral change: one involves ongoing action, or performance, and the other consists of learning, or acquisition (36, p. 478). Ongoing action reflects those properties of a person that describe his potentialities for action. Among these properties, according to Sears, are drives, habits, cognitive structures, and expectancies. The study of personality development requires a set of principles that will describe the way in which these properties, the potentialities for action, are changed by experiences. The changes in potentialities for action are best described by the laws of learning.

If Sears' concept of change is elaborated so that it explicitly includes behavioral manifestations that indicate shifts in performance and acquisition which are regressive, deteriorative, and declining in rate and proficiency, then it will prove to be a suitable frame-of-reference for the kind of developmental psychology called for by Dennis; a psychology adequate to the course of human life as a whole (4). The investigation of process, within this framework, includes utilization of geriatric and pathological samples, and the employment of laboratory techniques designed to impair as well as enhance performance.

A comparative investigation of ongoing action and acquisition raises questions not only concerned with the way in which the latter alters the former, but also is concerned with the influence exerted by the properties of ongoing action on the course of acquisition. A comparative approach necessarily emphasizes organismic characteristics, and centers interest on differences as well as similarities in the basic mechanisms of modifiability. The nature of the organism assumes prime importance in comparative investigation. That this is not always the case is amply demonstrated in Postman's operational definition of a perceptual experiment. The operations which define the perceptual experiment include (1) manipulation of the stimulus along selected physical dimensions, (2) certain specific manipulations of the organism and, (3) the measurement of a limited class of response (31). Comparative approaches add to this definition certain specifications, crude or otherwise, concerning the nature of the organism that provides the measured responses. Such definition is sometimes attempted by describing antecedent events. It is possible that description of this sort is ultimately adequate. In practice, however, it is likely to be quite incomplete; generally it is confined to laboratory manipulations, in which case it becomes synomymous with (1) and (2) above. Hilgard recognizes this problem directly when he asserts that paranoids do not use their past experiences the way other people do when they interpret what is going on in the present. He points out that gestaltlike principles may be necessary in accounting for how contemporary organization draws upon past experience in the achievement of perception (13). A similar argument is presented by Luchins who questions the assertion that all organization is based on past experience and sees the need for experimental inquiry into the principles governing the acquisition and operation of experience (22).

The potency of frequency, frequency of reinforcement, contiguity and like parameters for behavior is not denied when organizational factors are introduced. The function of such introduction is to direct experimental attention to certain divergencies between cognition and neo-association theories. A central problem posed by these divergent viewpoints involves the place, if any, in behavior theory, of what cognition theorists have called central representational processes. The adequacy of the classical Gestalt rendtion of these central representational systems or schema has been vigorously questioned from a number of directions. Piaget questions the Gestalt assertion of the permanence of certain forms of organization throughout mental development (29). Hebb discusses innately given, sensorially determined unities in perception and supplements these with nonsensory unities and identities primarily dependent upon experience (11). Harlow introduces experimental evidence which raises questions about the "equilibrium" assertions of Gestalt theory (10). Krech critically analyzes the sufficiency of the law of pregnanz for the description of behavioral events (19). This is but a sampling of the broad theoretical attack being levelled at orthodox Gestalt theory. This attack creates a number of problems for research. The attackers by no means rule out such things as figure-ground organization, progressive differentiation, or the notion of levels of functioning. Hebb retains *figure-ground* as an important conceptual tool, and in later writings (12) deals with motivation in a manner reminiscent of Lewin's discussion of valence (20). Krech amends the law of pregnanz to include the spontaneous development of states of greater heterogeneity and complexity, but his *Dynamic System* remains a field complex (19). Piaget retains and expands his position on the importance of schemata as determinants of behavior. In discussing the nature of the assimilatory act, he stresses the importance of differences in behavior patterns which characterize different periods of development (29, pp. 150-153). Assimilation is defined as the action of the organism on objects which surround it (29, pp. 4-8). There is a progressive differentiation of operations. Early operations, genetically speaking, are bound to immediacy, to that which is directly given in perception. More lately emerging operations permit conceptualizations which involve accommodation to contexts not directly given in perception (1, p. 30; 30, p. 189 ff).

Werner's system represents a similar modification of classic Gestalt theory. He describes development in terms of a series of organizational levels (43). Primordial behavior tends to be syncretic and diffuse. It is succeeded by gradually increasing differentiation and hierarchic integration. At both early and mature periods of development organization tends to be wholistic (5, 44). Primitive wholism is characteristically undifferentiated whereas the more mature whole manifests good articulation of parts and integration of parts into an encompassing system.

Murphy's discussions of development are akin to those of Piaget and Werner. Murphy stresses the sequential appearance of global, differentiating, and integrating function patterns, and points out that these patterns manifest themselves in growth, and in learning in both genetic and agenetic situations (27, pp. 66-68; 28, pp. 342-343).

Scheerer, also, suggests that representational modes active at various periods in development may retain their functional potency and reassert their influence under certain special sets of conditions in the adult (35, p. 9).

A conception not dissimilar in form to that offered by the cognitive writers is Hebb's description of the development of perceptual learning. He states:

> The course of perceptual learning in man is gradual, proceeding from a dominance of color, through a period of separate attention to each part of a figure, to a gradually arrived at identification of the whole as a whole: an apparently simultaneous instead of a serial apprehension (11, p. 32).

The central representational processes or schemata may be thought of in terms of Sears' "ongoing action or performance." For practical experimental purposes they must be considered as limiting factors in behavior. The refinement of certain orders of performance should, if such schemata are operative, be restricted along predictable lines. Systematic training series, Sears' "learning, or acquisition" variables, should produce differential learning effects in populations defined in terms of inferred organizational characteristics. Such population definition may proceed in terms of predetermined criteria, e.g., chronological age, intelligence, conditions of pathology, or in terms of performance in special test procedures utilizing tasks specifically designed to produce performance variations. Training procedures, for heuristic purposes, ought to be conducted along the lines suggested by systematic learning theories.

Tests employed to differentiate or categorize individuals with respect to organizational characteristics must be adequate to the task. Razran draws a distinction between "mere acquisition" of conditioned reactions in which perception is unnecessary, and acquisition under conditions involving perceptual relations. He states:

> Or in other words, whenever learning takes place, it involves the operation not of one kind *or* the other kind, but of either (*a*) the lower kind or (*b*) the higher *and* lower kinds. And specifically it obviously means that relational learning subsumes conditioning. . . . (32, p. 92).

Razran's discussion is pertinent to the question of the selection of test materials because the schemata or representational systems proposed by cognitive theorists are testable only when the task fulfills condition (*b*) above. Learning situations designed to subject the inferred system to test must involve the operation of relational factors. This point was clearly made some time ago by McGinnis in her study of the motor habits of young children (24, pp. 289-290). She found no relationship between first trial performance in a pattern maze and first trial preformance on a concealed-cue maze. While first trial performance on the pattern maze was related to intelligence, no such relationship was found between first trial performance on the concealed-cue maze and intelligence. McGinnis' pattern maze (the Young Slot Maze) is adequate for testing inferred central representational systems since first trial learning involves the discernment and utilization of path relationships. The concealed-cue maze is not an adequate testing device since first trial learning involves strictly rote procedures.

As an initial step in setting up a research program directed at exploring the relevance of cognitive theory for behavior, a series of base performance characteristics must be obtained on variously defined population samples. This is particularly necessary since cognitive theorists have tended to be relatively vague concerning a number of their constructs. A programmatic research example of the specification of organizational properties is provided by Werner (44), Friedman (5), and Siegel (38). Utilizing the Rorschach blots, and concentrating on structural features, a genetic scoring system has been developed. For example, two kinds of *W* responses are distinguished:

> . . . genetically low responses are those which indicate amorphous, diffuse, or confabulatory percepts where little attention is given to part relations and to perception of contours. The genetically high percepts are reflected in the responses whereby the percept is that of a precisely formed unit with parts integrated into it, where the whole is composed of relatively independent sub-wholes. . . . with age there is a decrease of the undifferentiated diffuse whole and detailed responses, there is an increase of highly articulated, well-integrated whole and detail responses. There is further interesting shift from the early whole responses toward small details between the ages of about 6 to 8 years declining in favor of the integrated whole response later on (44, p. 4).

Friedman reports that schizophrenics exhibit features of structural regression in the Rorschach which resemble the perceptual performance of normal children (5). Siegel, working with subjects characterized by degree of impairment, i.e., paranoid, and hebephrenic and catatonic schizophrenics, reports that the paranoid group produces Rorschach responses similar in structure to those obtained from normal children between the ages of six and ten years, while hebephrenic and catatonic subjects produce responses similar in structure to those found among normal children in the three to five year age group (38). Werner reports a series of "microgenetic" changes which resemble the ontogentic changes described above. When normal adults are presented with Rorschach blots under exposure times of 1/100 of a second, 1/10 of a second, 1 second, and 10 seconds respectively, the structures of the perceptual responses parallel those obtained from ontogenetic samples under conditions of prolonged exposure.

These Rorschach studies provide not only increased specification of cognitive inferences, but also suggest a methodology for integrating developmental, general, and clinical psychological problems within an experimental framework. This framework will prove fruitful if it is expanded to include parameters usually emphasized by systematic learning theories.

An experimental program along these lines is currently being launched at the Institute of Child Welfare at Minnesota.[1] It is concerned primarily with the development of visual and haptic recognition. Recognition occurs when an appropriate designating response is made by a subject to a stimulus with which he has had previous commerce. Such designating responses are not confined to exact stimulus reproductions, as shown by the extent of recognition behavior in everyday life. Individuals recognize many objects with which they have not had direct experience. Persons also frequently recognize objects which are partially obscured, disguised or altered with respect to their original representation. E. J. Gibson's review of pertinent literature presents investigations which cover impoverishment of stimuli, effects of practice and reinforcement, and transfer. She concludes that there is ample evidence that perceptual learning occurs under many conditions (6).

Despite the vast array of data on conditions under which discrimination and recognition are enhanced, we know little concerning the development of these behaviors. What are the governing principles? Association theorists propose that behavior in general, including recognition, develops gradually and is subserved by reinforcement (14), contiguity (9), redundancy (26), etc. Cognition theorists introduce the notion of saltation in development and assert that behavior occurs as a function of the underlying patterning of psychological processes. Differentiation of psychological function imposes limits on the nature of performance, and behavior is accounted for in terms of the inferred degree of differentiation (18, 20). More extensive descriptions of developmental changes in central patterning are provided by Murphy (27), Piaget (29, 30), and Werner (43). The assertions contained in these alternate points of view suggest an experimental series in the area of visual and haptic (stereognostic) recognition which may serve to clarify particular theoretical issues.

---

[1] This research is aided by a grant from the National Science Foundation.

Initially it is important to select experimental tasks that contain selected features and are capable of eliciting wide arrays of individual performance differences. This procedure is accompanied by the selection of subjects in terms of (cognitive) theoretically-inferred characteristics. Characteristics of developmental samples have been specified both by cognition theorists, and others (25). The selected test materials should elicit performance differences consistent with the predictions made about the several subject groups. Once such base performance rates are available specific tests can be made concerning the behavioral effects of the parameters of systematic learning theories. Subjects are drawn from virtually the entire age range, represent various levels of intelligence, and degrees and conditions of pathology. This basic procedure is to be replicated with a number of tasks and in several perceptual modalities.

A pilot study exemplifying the initial phase of basic procedure has been completed. Subjects in this study include 48 preschoolers from the Nursery School and Kindergarten of the Institute of Child Welfare and 17 adults drawn from graduate courses. The materials consist of 15 common objects, e.g., umbrella, automobile, ice cream cone, etc. These objects are presented on cards, one at a time, and identification is requested. The contours of the pictures are produced by a series of small cross-hatches and obscured or embedded by crosshatching over the face of the card. After familiarization trials, *S*s of preschool age received exposures of four seconds duration while for adults the exposure time was one second. Table 1 presents the mean number of correct identifications for the several age groups.

TABLE 1

MEAN NUMBER OF CORRECT IDENTIFICATIONS OF MEANINGFUL EMBEDDED FIGURES BY AGE GROUPS*

Age	*N*	*Mean*
2-6 — 3-5	17	1.00
3-6 — 4-5	18	2.78
4-6 — 5-5	13	5.23
Adults	17	11.53

* Ages 2-6 to 5-5 received four-second exposures while adults received exposures of one second.

The beginnings of a base performance rate for the particularly characterized groups begins to emerge. Alternate sets of explanations for these performance differences are available from association and cognition theories. The latter place emphasis on differential figure-ground facility while the former stress such aspects as frequency of prior exposure, or reinforcement, etc. Systematic training series with these materials is currently being undertaken. The effects of training will reflect directly on the various theoretical assertions which have been reviewed.

Another experiment, utilizing different materials, has provided similar ranges of performance variation among preschool populations (7). Once again common objects were used as stimuli. Each object was represented five times to constitute a series involving a gradation of completeness. The initial card contains very few contour lines or internal details and these are gradually built up through the series until on the final card the complete object is present. There are eleven such series of five cards each. Only *Ss* who have the appropriate verbal designating response for the complete pictures are included in the analysis. Even under these conditions there are marked differences among age groups in terms of the degree of completeness of representation required before recognition occurs.

A similar line of investigation is being carried forward with geometric figures under conditions of embeddedness. Pilot results confirm not only that marked performance differences occur within age groups as reported elsewhere (8, 49), but also between age groups. There is some suggestion that decline in performance sets in relatively early in middle age.

Keller's work with a mentally retarded boy suggests another experimental procedure involving not only recognition but also reproduction (16). Keller used Kohs block designs and requested his subject to reproduce the designs on paper which contained squares to represent the blocks. The stimulus was left before *S* while he made his copy. No time limit was imposed. *S* failed in his reproductive attempts, evidently being unable to come to the realization that the stimulus could be accurately reproduced on the squared paper by breaking the stimulus up into imaginary squares. In order to rule out the possibility that *S* was unable to discriminate between stimulus and reproduction he was presented with the test figure for five seconds and then required to select it from among four other figures, one of which was the correct match, and another of which was his own previous reproduction. Under these conditions *S* quickly and accurately selects the correct figure. *S* was next required to reproduce a test figure which was presented on squared paper identical with that used for the reproductions. *S*, despite this perceptual crutch, still fails. Initially, he was occasionally successful but as he proceeded with a design his reproductions suffered. Keller accounts for this behavior in terms of the "tyranny of the figure-as-a-whole." When *S* is required to reproduce diagonals with minimal configural properties he succeeds. As soon as the diagonals are joined in a configuration, *S* fails. In a final experiment *S* was provided with a frame which masked all but one square of the Kohs stimulus, and he was taught to go from square to square, building the design piecemeal. Keller reports:

> It was thought that this crutch might teach him that the total configuration could be periodically ignored and that the design could be built solely

> by careful attention to the individual units of which it was composed. He was able to operate very successfully with the frame, but when the frame was removed he reverted to his old confusions. Apparently, he could not by conscious control focus successfully on the parts so long as the whole was there to distract him (16, p. 564).

These findings are more striking when it is recognized that the entire procedure was conducted with therapeutic intent and every effort was made to help *S* improve his performance. These results reflect neither a simple perceptual defect nor a motor defect. The defect has been labelled a *specific cognitive defect* by Keller.

This kind of "functional analysis" is characteristic of much of the work of Kurt Goldstein and of the "brain-injury" researches of Heinz Werner. The specific techniques utilized by Keller are well suited for investigations designed to elicit performance differences in developmental and clinical samples and to determine the extent to which these differences are susceptible to modification by systematic training. The nature of such techniques permits inferential construction of central dynamisms which have predictive power.

The visual recognition techniques of Lowenfeld (21) with the modifications introduced by Wiggin open up still another path of potent exploration. Lowenfeld pulled line drawings behind a slot to that *S*s perceived only a limited aspect of the total configuration at any moment in time. *S* was then required to locate the figure among a group of figures presented for inspection. Wiggin standardized presentation by recording the stimuli on motion picture films (48). These techniques are applicable with both meaningful and nonsense materials.

For the entire series of visual recognition and reproduction experiments such factors as masking, interruption, time of exposure, context, etc., will be manipulated. The research program will supplement techniques previously discussed with those utilized by Werner, and Werner and Strauss with brain-injured children (39, 42, 45, 46, 47). These techniques address themselves nicely to the central research problem; namely, the specification of inferred central representational processes and the behavioral consequences of these processes for learning.

The basic design for research on learning, recognition, and reproduction in the visual modality is also applicable in the relatively unexplored field of stereognosis. A limited number of studies with animals (e.g., 34), in industrial psychology (15), in psychopathology (47), and in the area of creative activity have been conducted (21). These investigations contain promising techniques which can be modified to fit present needs.

Lowenfeld's work with tactual materials suggests a number of techniques useful in obtaining base performance rates on developmental samples. He presents *S*s with solid objects for tactual exploration and requires them to identify the stimulus object from among a group of objects. He also used stencils cut in heavy cardboard. Woodworth reports a study by Yamane in which braille figures were utilized with blind and sighted-blindfolded *S*s (50). Ruch, Fulton, and German used solid objects with primates under conditions of parietal intactness and parietal destruction (34). Werner and Strauss presented brain-injured *S*s with relief figures and required their *S*s to reproduce the stimulus (47). W. O. Jenkins employed various shapes and shape orientations in the course of developing a system of easily discriminable aircraft controls (15).

Revesz found rules for haptic perception akin to those obtaining in visual perception (33). There are very few developmental data on hand for this perceptual modality: collection of data will add information as well as serve theoretical ends. Modifications of available techniques will be utilized to obtain base performances and to test the assertions of association and cognition theories with respect to the specified variables as indicated by the basic design.

The haptic materials can be embedded, interrupted, varied in time, meaningful or nonsensical, varied in texture, pliability, and in size. Stereognosis may be an important area for exploration if its investigation adds essential data relevant to theorizing about learning and perception.

An adequate comparative psychology must deal not only with interspecies similarities and differences but also with intraspecies similarities and differences. General behavior theory has much to gain from the systematic collection of data within a developmental-comparative framework. These investigations may be fruitfully conducted with the aid of descriptive schemata and response inferred mediational mechanisms akin to those proposed in this paper. The value of such conceptions for biology is demonstrated by the notion of "organizing factors" introduced by the embryologist Paul Weiss (41), and more recently by the ontogenetic analogue of the course of nonspecific stress reactions in disease presented by Hans Selye (37).

In sum, this orientation permits specification of the properties of ongoing action and the relationship between ongoing action and acquisition variables. Predictions concerning the effects of inferred ongoing action properties are drawn from cognition theories. Since these predictions often are at variance with those of systematic learning theories, some clarification of this divergency is likely to emerge.

REFERENCES

1. AEBLI, H. *The development of intelligence in the child: summary of the works of Jean Piaget between 1936 and 1948.* Minneapolis: Institute of Child Welfare, Univer. of Minnesota, 1950.
2. BARKER, R. G. Child psychology. *Annu. Rev. Psychol.*, 1951, 2, 1-22.

3. Beach, F. A. The snark was a boojum. *Amer. Psychologist,* 1950, 5, 115-124.

4. Dennis, W. Developmental theories. *Current trends in psychological theory.* Pittsburgh: Univer. of Pittsburgh Press, 1951. Pp. 1-20.

5. Friedman, H. Perceptual regression in schizophrenia: an hypothesis suggested by the use of the Rorschach. *J. genet. Psychol.,* 1952, 81, 63-98.

6. Gibson, Eleanor J. Improvement in perceptual judgments as a function of controlled practice or training. *Psychol. Bull.,* 1953, 50, 401-431.

7. Gollin, E. S. Developmental differences in visual recognition under conditions of stimulus impoverishment. In preparation.

8. Gollin, E. S., & Baron, A. Response consistency in perception and retention. *J. exp. Psychol.,* 1954, 47, 259-262.

9. Guthrie, E. R. *The psychology of learning.* New York: Harper, 1935.

10. Harlow, H. F. Learning theories. *Current trends in psychological theory.* Pittsburgh: Univer. of Pittsburgh Press, 1951.

11. Hebb, D. O. *The organization of behavior: a neuropsychological theory.* New York: Wiley, 1949.

12. Hebb, D. O. Drives and the c.n.s. (conceptual nervous system). *Psychol. Rev.,* 1955, 62, 243-254.

13. Hilgard, E. R. The role of learning in perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality.* New York: Ronald Press, 1951. Pp. 95-120.

14. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.

15. Jenkins, W. O. Tactual descrimination of shapes for coding aircraft-type controls. In P. M. Fitts (Ed.), *Psychological research on equipment design.* Washington, D.C.: U. S. Government Printing Office, 1947.

16. Keller, J. E. The investigation of a specific cognitive deficiency. *Amer. J. ment. Def.,* 1954, 58, 560-565.

17. Koch, H. L. Child psychology. *Annu. Rev. Psychol.,* 1954, 5, 1-26.

18. Koffka, K. *The growth of the mind.* New York: Harcourt, Brace, 1925.

19. Krech, D. Dynamic systems as open neurological systems. *Psychol. Rev.,* 1950, 57, 345-361.

20. Lewin, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.

21. Lowenfeld, V. Tests for visual and haptical attitudes. *Am. J. Psychol.,* 1945, 58, 100-111.

22. Luchins, A. S. An evaluation of some current criticisms of gestalt psychological work on perception. *Psychol. Rev.,* 1951, 58, 69-95.

23. McCandless, B. R. Special review: recent books in child psychology. *Psychol. Bull.,* 1955, 52, 150-158.

24. McGinnis, E. The acquisition and interference of motor habits in young children. *Genet. Psychol. Monogr.,* 1929, 6, 203-311.

25. Miles, W. R. Age and human ability. *Psychol. Rev.,* 1933, 40, 387-414.

26. Miller, G. A. *Language and communication.* New York: McGraw-Hill, 1951.

27. Murphy, G. *Personality: a biosocial approach to origins and structure.* New York: Harper, 1947.

28. Murphy, G., & Hochberg, J. Perceptual development: some tentative hypotheses. *Psychol. Rev.,* 1951, 38, 332-349.

29. Piaget, J. *The psychology of intelligence.* New York: Harcourt, Brace, 1950.

30. Piaget, J. *The origins of intelligence in children.* New York: International Universities Press, 1952.

31. Postman, L. The experimental analysis of motivational factors in perception. In *Current theory and research in motivation: a symposium.* Lincoln: Univer. of Nebraska Press, 1953. Pp. 59-108.

32. Razran, G. Conditioning and perception. *Psychol. Rev.,* 1955, 62, 83-95.

33. Revesz, G. *Psychology and art of the blind.* New York: Longmans, Green, 1950.

34. Ruch, T. C., Fulton, J. F., & German, W. J. Sensory discrimination in monkey, chimpanzee, and man after lesions of the parietal lobe. *Arch. Neurol. Psychiat.,* 1938, 39, 919-937.

35. Scheerer, M. Personality functioning and cognitive psychology. *J. Pers.,* 1953, 22, 1-16.

36. Sears, R. R. A theoretical framework for personality and social behavior. *Amer. Psychologist,* 1951, 6, 476-483.

37. Selye, H. Stress and disease. *Science,* 1955, 122, 625-631.

38. Siegel, E. L. Genetic parallels of perceptual structuralization in paranoid schizophrenia: analysis by means of the Rorschach technique. *J. proj. Tech.,* 1953, 17, 151-161.

39. Strauss, A. A., & Werner, H. Disorders of conceptual thinking in the brain-injured child. *J. nerv. ment. Dis.*, 1942, 96, 152-172.

40. Verplanck, W. S. Since learned behavior is innate, and vice versa, what now? *Psychol. Rev.*, 1955, 62, 139-144.

41. Weiss, P. *Principles of development: a textbook in experimental embryology.* New York: Holt, 1939.

42. Werner, H. Development of visuo-motor performance on the marble board tests in mentally retarded children. *J. genet. Psychol.*, 1944, 64, 269-279.

43. Werner, H. *Comparative psychology of mental development.* New York: Follett, 1948.

44. Werner, H. Developmental approach to general and clinical psychology. Presented as part of a symposium at Massachusetts Psychol. Ass., March, 1954.

45. Werner, H., & Strauss, A. A. Types of visuo-motor activity and their relation to low and high performance ages. *Proc. Amer. Ass. ment. Def.*, 1939, 44, 163-168.

46. Werner, H., & Strauss, A. A. Causal factors in low performance. *Amer. J. ment. Def.*, 1940, 45, 213-218.

47. Werner, H., & Strauss, A. A. Pathology of figure-background relations in the child. *J. abnorm. soc. Psychol.*, 1941, 36, 230-248.

48. Wiggin, R. G. A criticism of two of Lowenfeld's visual-haptic tests. Unpublished paper, Institute of Child Welfare, Univer. of Minnesota.

49. Witkin, H. Individual differences in ease of perception of embedded figures. *J. Pers.*, 1950, 19, 1-15.

50. Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* (Rev. Ed.) New York: Holt, 1954.

THE REPLACEMENT PROBLEM IN CHILD DEVELOPMENT

Harold E. Jones
University of California

Three decades have passed since the Laura Spelman Memorial began its program of subsidies to child development. The establishment of institutes, a generous fellowship program, and financial aid to journals produced a prompt increase in the number of persons annually entering the field, and correlatively a sharp rise in published research.

A part of this growth was normal, to be expected from the general growth of psychology and also from the earlier stimulus of the child study movement. A substantial part of it, however, was an outcome of forced feeding, and it may now be appropriate to inquire if this nutritional experiment was a sound and wise one.

The answer to this question may be a qualified affirmative—qualified in view of certain side effects which have not been wholly favorable to the normal growth of the field. One such effect was the almost inevitable counter-reaction of some persons at the core of the older disciplines, who felt that too much attention, and too much money, were being given to a new and untested branch of scientific effort.

In some psychology departments, steps were taken which to a degree weighed against graduate students with an interest in this field, and which were later to be used even more vigorously as a defense measure in dealing with a flood of clinical students. These sanctions sometimes took the form of graduate requirements heavily emphasizing historical and systematic psychology, and laboratory method. The prestige fields of psychology, closest in spirit to natural science, have always had some success in repressing offshoots. I do not wish to imply that there is anything malignant about this. Perhaps we should regard it chiefly as a kind of homeostatic control within the science, necessary to keep things in balance. Child development,

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

once in a stage of perhaps too rapid growth, encountered inhibitors, and some of these inhibitory mechanisms are still effective.

Another factor which needs to be considered is that the initial program of recruitment around 1930 brought into child development a number of persons who already had training and loyalties in other fields, marginal to science. It must be admitted that their interest in children did not necessarily imply a devotion to scientific method.

In the middle and late 1930's the artificial stimulation began to subside. A number of the centers originally dependent on foundations were undergoing a transition to local support. A fairly good balance was established between supply and demand of research workers and teachers in child development, and it was reasonable to expect growth—no longer at a pace of positive acceleration but of a slow and orderly nature.

World War II, however, brought new changes which were for the time being unfavorable to child development research. This field had little to offer in the prosecution of the war effort, or in dealing with veterans' problems after the war. In the national interest, very large resources were devoted to recruiting psychologists into clinical work, and also into various aspects of military psychology. Of all the branches of psychology, Developmental has undoubtedly suffered the most from the subsidizing of clinical training. Many graduate students who would otherwise have worked in child research, and whose interests were not primarily in therapy, could hardly be expected to resist the stipends and the superior resources for training which were made possible, on a grand scale, by governmental agencies.

Since 1952 somewhat greater funds have become available for fellowships and grants-in-aid in child development. The National Institute of Mental Health has had a liberal policy in this regard, and we now know that many aspects of child development, especially on the biological side, are relevant to the program of the National Science Foundation.

Although no foundation has taken the place of the Laura Spelman Memorial, in a primary concern with children, we can no longer feel that the sun is under a cloud so far as research support is concerned. There remain however many problems in our recruiting program. From the standpoint of the recruit, what kind of a field will he or she be entering? Is it one which presents a cheerful front to the world, with some pride in the past and confidence in the future, or is it one marked by self-doubts and querulous complaints?

After a recent visit to this country, Dr. Skard of the University of Oslo reported on Recent Trends in Child Psychology in U.S.A., in *Courrier*. Her opening statement is that, "The most striking impression that research in child psychology gives today is one of great differences and divergencies in problems and methods. A latent dissatisfaction seems to be present among the scientists in this field." Now, this is much better than accusing us of complacency. But dissatisfaction may be of several kinds. It may be constructive or it may be destructive. A lively interchange of criticism of specific studies may help to improve methods and lead to new answers to new problems. But if there is a general melancholia about the field, or a cynical attitude about all approaches except one's own, we can hardly expect to present an attractive picture to students.

Instead of deploring divergencies, I think we should regard individual differences in the *modus operandi* of research as a very favorable sign. We should recognize that psychology, and child development, flourish on individual differences. It is, I think, extremely fortunate that we have vigorous interests in laboratory research on theoretically loaded problems. It is also of great importance for the future of the science that we have an expanding interest in theoretically neutral research, and in the patient application of exploratory observational methods. There is a bright promise in cross-cultural research, and in the renewed application of comparative or cross-species methods in studying some genetic problems. We should welcome the fact that we can shelter and provide expression for a great variety of emphases in clinical, quantitative, descriptive, interpretive, naturalistic, experimental, and longitudinal approaches.

But it is a somewhat puzzling fact that a few persons who are happy in only a limited corner of the field seem unhappy that other areas exist. Their efforts to devalue what they are not themselves doing may result merely in devaluing the whole field. This situation has been examined earlier in this conference, with regard to the strategy of research, but my concern with it here is in connection with the effect upon recruitment.

Of course, when we consider the actual number of persons, among psychologists, who are working with children we find that our clientele is by no means a vanishing one. In 1918 there were only 3 psychologists who indicated a primary interest in children. By 1937 this number had increased to about 81. At the present time nearly a thousand members and fellows of the APA indicate the child, or work with children, as their first interest. It is however true that too few of these are active in research, and that measures are needed to increase the training of research personnel.

Proposed measures may be considered first at the national level and then in terms of what might be done in individual institutions. Obviously the most important factor in recruiting students is to have something to recruit them into: an active and healthy research field, and one which is not too much under the influence of the need for aggression or the need for abasement. In order to compete with other fields which are already well supplied with stipends, I would propose a national program of competitive fellowships specifically for child development research. We might appropriately discuss how such a program could be set up, and under what auspices. Even if it included no more than five or ten awards a year, it would have a value going considerably beyond the actual number of appointments. I would also advocate a national committee to deal with problems of child development teaching, to administer visiting appointments, and to study instructional programs, instructional aids, and graduate training curricula. This cannot very well be done by an *ad hoc* committee appointed by a society. It should have funds, and representation from different societies.

Finally, the effect of such measures should be to extend the number of institutions which, like Minnesota, Chicago and Iowa, have a training program specifically in human growth and development. Where a student interested in child development must take his degree in a psychology department, he sometimes finds that the demands of the "classical revival" make it all but impossible for him to conduct a well-rounded program. He is well trained in psychological systems and conventional laboratory methods, but finds it difficult to get what he needs in physical growth, developmental physiology, or in courses in sociology or anthropology bearing on child development. Lacking time to conduct a field study, he may find that the line of least resistance toward a Ph.D. is through the rat laboratory or in the customary exploitation of college sophomores.

Every now and then I like to refer back to a statement made more than a quarter century ago by W. I. Thomas, in reference to the genetic method, which as he said represented "a reaction against the practice of observing behavior at a given moment by a given technique (e.g. psychological, observational) and drawing inferences as to causation. It is realized that the conception can no longer be entertained that it is adequate to correlate any two variables—a given behavior manifestation with nutrition, or skeletal development, or mentality, or birth trauma, or family, or bad company, taken separately. The genetic method proposes, therefore, to study human growth as a maturational and experiential process from early infancy through adolescence and beyond, in individuals, groups and control groups, with reference to both physical maturation and the complex of cultural factors (family, school, moral code, etc.), by the simultaneous and continuous employment of all available techniques—anatomical, physiological, psychological, etc."

For a generation ago, and perhaps for the present generation also, this may be too ambitious a statement. Our child development research must often be contented to whittle away at some parts of this interdisciplinary programming. We must make many compromises. But we should make them reluctantly rather than triumphantly, and we can hardly forget that if we are really interested in child development we must sooner or later come to grips with changes through time, and not merely with the horizontal cross section. Graduate programs which are aimed only at cross section or only at situational research may not be well adapted to recruiting and training all of the kinds of persons needed in this field.

THE NEED FOR CONCEPTUALIZATION IN RESEARCH ON CHILD DEVELOPMENT

IRVING E. SIGEL
Merrill-Palmer School

To be asked to pose research priorities in child development presupposes that we are dealing with a contained discipline in which basic issues and problems are clear. Child development, however, at this time does not have a coherence, and consequently a statement of research priorities emanates from idiosyncratic perceptions of the field rather than from deductions derived from the conceptual unity of the field.

In order, then, to discuss priorities, it seems that we must come to grips with a prior question, namely, what is child development? The answer to this question is a necessary step in posing priorities in research, for ideally they can flow from the problems residing within the area.

The question, what is child development, is old. Today, it has various connotations. To some it is merely an interest area; to others it is a point of view; to others it is an area with a specific content; to still others it is an empirical hodgepodge of unrelated facts and speculations dealing with the child.

For me, child development is an aspect of human development. The emphasis on the childhood period is partly idiosyncratic and partly based on the assumption that the basic patterns of many human characteristics and processes have their beginnings here. The goal then of child development is to understand the processes and content of the "progressive" change of the organism, specifically differentiation, specialization and integration of human structures and functions as the organism moves along the continuum of the life cycle. The functional unity of child development and of its broader aspect, human development, stems from the common integrating principle of development. The developmental set is crucial in distinguishing child development from other positions studying the child.

If the premise is accepted, then some of the basic research questions facing us, derived from the developmental orientation, are as follows: what changes occur within the organism as he matures, what conditions affect the course of this change, what processes are involved in change, and how these changes influence and are influenced in the organism's interacting and coping with his environment.

With the task so defined, the question becomes, how do we go about it? The purpose of this paper is to propose a procedure for so doing in order to facilitate the achievement of the goals stated above.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

If a segment of human knowledge is to be organized under a single rubric for the purposes of scientific study, it behooves us to work towards the establishment of a conceptual framework within which a theory or theories can be evolved. Before proceeding with such a comprehensive task, we should ask ourselves if there is any current conceptual orientation which attempts to integrate knowledge from various established disciplines into a comprehensive unit.

The "holistic approach" is one point of view which is prevalent among child development workers. It holds that we are concerned with the total organism, that the organism in its unity is different from the sum of its parts. If this assumption is correct, the course of action to be followed should stem with methodological consistency from such a conceptual position. It requires that methodologies be set up so that the whole can be studied simultaneously in all its complexity. To my knowledge this has never been done, nor at this point in our development, do I believe it can be done.

Whether the organism is such a totality functioning in such unity is still an open question. The acceptance then of the assumption precludes the investigation of relationships of aspects of the organism, since this, too, does violence to the essential nature of the individual.

Further, this position leads to the conclusion that our existing knowledge about human development is in error, since little if any of it was gathered under holistic conditions. The organism has been studied at different time intervals and in its different functions. Yet we know that we have bodies of verifiable knowledge useful in predictions of its future reactions.

Thus I see this holistic orientation as being inconsistent, at least in regard to theory and practice. For a unified science we need theory consistent with practice. This then means that the empirical research within a discipline should ideally be compatible with the theoretical position or consistent with the conceptual framework taken.

The holistic position, to me, is a philosophical one in which highly abstract concepts can be used to integrate the totality of man in his universe. But to presume that a body of empirical knowledge can emanate under such a point of view seems highly doubtful. If we leave the holistic approach to the realm of philosophy, I believe we will be in a better position in evolving a conceptual framework which can relate coherently to empirical research.

Rather, I would prefer to think of the "whole" child as an organization of a number of systems, which function at varying degrees of autonomy and interrelatedness.

The first step in organizing such a conceptual framework is the identification and delineation of systems within which we can organize coherent developmental material. Next we need to gather together existing data describing the course of development within the designated systems. After this, we should aim to describe the stages of development within systems and between systems in order to describe the concomitant appearance of structures and functions; with this achievement we can proceed to examine relationships between systems to determine the degree, the kind, and the quality of relationships existing between systems.[1,2]

In setting about to delineate and define systems of the organism we become impressed with the unalterable fact that the course of a child's development is a function of a variety of variables residing primarily within the province of biological and social sciences. Thus the science of development seeks relevant variables in such disciplines as biology, psychology, anthropology, to list but some. The need to be dependent upon many disciplines in our construction of a single conceptual framework stems naturally from the nature of our problem.

We are immediately faced with the question, *what* do we incorporate from the various disciplines? Do we take concepts, or facts? Do we take theories? Do we take whatever we deem necessary? If so, what [illegible] criteria for this necessity?

Depending on whether we incorporate theory, concepts, assumpti[illegible] facts, we face different problems. When we borrow concepts we must a[illegible] tain their meaning and their utility for answering our own needs. Is t[illegible] concept borrowed relevant for our material? For example, homeostatsis is a case in point. It is used in physiology and in psychology. Is its meaning equivalent in both instances? Or are we just borrowing words, which we then redefine to meet specific conceptual questions? Or are we attempting to fit a concept onto a set of data which has been derived from very different sources? I feel these questions must be posed when we attempt integration of any kind. The flight into someones else's conceptual structure may be more impeding than facilitating.

We can ask similar questions when we attempt to incorporate theoretical positions. Is it a flight into theory because that theory has direct meaningful relationships and utility for our purposes? Or is it rather an expediency? In other words, do we try to set the theory onto a given set of empirical facts or does the theory emanate and develop from various observations and interpretations of observed phenomena? Again relevance, utility and appropriateness should be used as guides. One more point here is that we should also examine the assumptions upon which a theory is based. Such assumptions, altogether too frequently are not explicitly stated, and may be inherently incompatible with that of the borrower.

When we seek to incorporate facts from other disciplines, we also need to be cautious. The usual criteria employed in assessing any scientific fact

---

[1] To be sure, some steps have been taken already in the directions cited above. Baldwin's (1) recent text is a notable example of the application of psychological concepts to development. More concentrated efforts should be expended in these directions in the next few years.

[2] Although these stages are set in sequence, there is great overlapping. Thus, investigation of development within a designated system may also yield data on relations between systems. Thus the order here appears more rigid than intended.

should be employed. A fact obtained, however, under one conceptual system may not be a fact under another system. For example, if the corroborated findings of experimental research indicate that children from authoritarian settings are more punitive than children from nonauthoritarian backgrounds, we must be clear as to what punitiveness is. Facts then will have to be evaluated from the context of their origin.

Assimilating concepts, facts or theories from other disciplines does not mean, however, that child development has to be a receptive science—receiving contributions from others only. In the initial stages of organization and integration such receptive dependence may be necessary. As child development progresses in its unity, it will evolve more and more its own body of theory and fact, and hence move toward greater autonomy. However, it should be recognized that some interdependence of all the sciences exists. Let is not be thought that I conceive of child development as progressing toward insulation or isolation. Rather I conceive of it as developing in its unique way, facing its unique problems and so contributing to the study of man. In so doing it can furnish other disciplines concepts, theories and facts.

With some of these considerations in mind, the task facing us is to delineate the systems and define their content. Admittedly such a task is somewhat arbitrary, but for the purposes of study and research this approach is necessary. Let it be remembered, however, that the arbitrariness with which this is done is utilitarian and relative. This is no plea for absolutism.

We start with the premise that the developing organism in its environment comprises our field of interest. How we choose to delineate the systems depends on the level of abstraction at which we elect to start and the nature of the concepts we elect as guides. For example, we can start with such distinctions as biological, psychological and social. Or we can start with the personality and the environment. Or we can proceed by viewing the total in a functional-structural dichotomy and proceed from there. The particular starting point derives from our basic conception of human nature and of our perception of the organization of knowledge.

A starting point for now may be a tripartite breakdown of biological, psychological and exogenous environmental systems. Within each of these further delineation is possible, e.g., within the biological we can separate skeletal, or neural subsystems; within the psychological such subsystems as cognition, perception can be distinguished. Further identification of smaller units within some of these second order categories can be isolated.

Hopefully, the concern with each of these systems would be with process and with content. Thus, the interest, in psychological systems, for example, would be with the processes, e.g., cognitive process, and with the cognitive content, e.g., attitudes, beliefs.

I should like to digress here momentarily to emphasize one of the three large systems mentioned above—the exogenous environment—since the consideration of the exogenous environment as a functional variable has received little systematic attention. In view of this rather serious omission, I should prefer to discuss this at some length, fully aware that similar elaborations are necessary for other factors.

As Chein (3) has pointed out, the environment must be systematically schematized so that appropriate representation of variables within the environment can be differentiated. The environment can be analyzed structurally, e.g., social structure, social class, family organization. What relationship exists between the child and various aspects of his environment? If we conceive of the environment as geo-behavioral, that is, the properties of the physical-personal environment as responded to by the reacting individual, then we can conceptualize the environment in functional terms, e.g., supporting, constraining, or directing the development of the individual.[3]

From this point of view, then, we can investigate the differential influence of various environmental experiences upon the developing child. As the child matures, certain aspects of the environment may function differentially; e.g., the home environment, considered globally for illustrative purposes, may have a high frequency of supportive and directive points of contact during infancy, whereas in later childhood this may shift to greater constraint and less support.

Such emphasis on the environment is especially important in the study of the socialization process. Extensive description of various points of environmental contact would give us greater insight into the manifold experiences which are involved in the child's socialization. Barker and Wright (2) have taken some significant steps in this direction and it is necessary to extend such investigations into various other types of communities. This is especially true for children living in urban areas.

Once we have determined the relevant developmental systems we need to describe adequately the developmental course for each of them. What are the developmental trends in various aspects of cognitive development? Piaget (5), for example, has given detailed description of some of the stages of development in this area. Gesell (4), Shirley (6) and others have presented descriptive material on physical growth trends. Additional descriptive data are needed. Where the need is can be most readily ascertained once we begin some sort of organization of the separate systems and assess the state of knowledge within them. This procedure is based on the premise that adequate description of events is a necessary phase of growth in any science.

To be sure, we are not content with descriptive data, for we do not conceive of our science primarily in taxonomic terms. What we seek are explanatory principles and theoretical constructs which will further our understanding of developmental phenomena.

---

[3] The terms are those used by Chein, but the definition of the environment is not identical to his.

In the search for explanations and for laws of development, we are forced to examine the relationships between systems as well as the presumed homogeneity within systems. To what degree then is a specific system homogeneous? And further, to what degree is a specific system related to other systems?

As we proceed to answer these questions, we may arrive at clusters of systems which are presumably more interdependent than others. These clusters I like to call self-contained units. Let me illustrate. Take three systems: social structure, adult personality, and child rearing practice. We can examine each of these separately, if we seek answers to such questions as, what kind of social structure exists in a certain society, how does this society function, etc. Or if we are interested in personality or child rearing in and of themselves we can investigate each of these as a separate.

If, on the other hand, we choose to study how social structure influences adult personality,. we may arrive at the notion that it does so primarily through its child rearing practice. Hence we can examine the relationship among these three—a self-contained unit—without involving any other system. The nature of the questions posed, then, becomes one of the determinants of the organization of self-contained units.

Yet it may well be that there exist "natural" self-contained units, within which systems are more intimately tied. Our empirical research, testing our conceptual organization, will indicate the relative closeness of systems.

A further step is to conceptualize and test the relationship between self-contained units. The implication here is that even self-contained units may not be completely autonomous but may have only relative autonomy. By research derived from the type of conceptualization proposed here, we can determine the degree to which self-contained units are closed and have no relevant dependence, under certain conditions and to limited degrees, on other units.

It should be said here that the concept of self-contained units is not synonymous with the concept of a system. A system is an assemblage of processes or structures collectively contributing toward specific functions or processes. Thus, processes and structure of cognition, for example, can be viewed as a system. A self-contained unit may contain one or more systems. The criterion of belongingness is the degree of mutual interdependence. Thus, physical growth can be viewed as a self-contained unit comprising numerous systems, e.g., skeletal, neural, circulatory, etc. The concept self-contained units is heuristic at this time. In time, as our conceptualization improves, such heuristic concepts may prove to have theoretical value.

As soon as we begin to examine the interrelatedness of systems and/or self-contained units, particularly if these are derived from different disciplines, we are faced with problems of conceptual and factual comparability and equivalence. In the social structure—modal adult personality—child rearing question, for example, we seem to be dealing with extensive and abstractive concepts which appear equivalent. Thus, we would separate out aspects of each of these and indicate the relationships so that we could see what elements of social structure affect what elements of personality. Further, we would have to find out what aspects of personality affect what child rearing practices; and finally, what differential effects various child rearing practices have on the developing personality of the child. In this way, presumably lawful relationships can be empirically demonstrated so that our abstract generalizations about the social structure—adult modal personality—child rearing practices chain reaction can be made with some degree of specificity.

If, on the other hand, we do not tend to see conceptual equivalents and the variables differ in their level of complexity, without considering these, then we do not know what affects what. For example, when we seek relationships between physical and mental characteristics, we find low relationships. This may be due to the fact that the variables contain different degrees of complexity so that the correlations are low, not because *no* relationship exists, but rather elements within the variables not identified or made explicit are included which operate to reduce the relatedness. IQ and skeletal growth studies are good illustrations of this. The various complexities represented in the single IQ score may mask or counterbalance each other, or be unrelated. Thus, low or insignificant relationships are found.

A further problem that occurs here from the developmental framework is that at various stages different relationships between variables may exist. For example, in the first year of life, motor skill and mental ability are more closely interdependent, and seemingly causally interrelated, than at age ten. The infant needs certain motor skills to handle certain intellectual tasks, whereas at an older age, with language, motor skills assume a different role.

Thus, when we view the relationships between systems of the organism developmentally, we may need to indicate the relationships sequentially as well as concomitantly. Some systems, such as motor skill or physical growth, have a stage setting function—they prepare the organism for performing certain tasks, while other systems may be correlated in their appearance in time, such as fine motor skills and the beginnings of conceptual thought.

The schema presented for systematization of the material within the child development area is admittedly broad and extensive. The task of ordering the material into meaningful units requires a more detailed and specific conceptual frame-work than was described here. This I feel is a research task and is one which hopefully can proceed in the next few years.

If such a goal is deemed important, then it might well act as a guide to our selection of problems, developing of concepts and interpretation of empirical fact. As we proceed, then, in such a direction, we continue to strengthen the scientific basis for the field.

One practical approach to such systematization can be in our graduate teaching program. Graduate seminars can be organized around conceptual problems where efforts are directed toward definition and delineation of systems, integrating existing information within these systems and possibly developing concepts for hypothesizing relationships within and between systems. Hopefully, a task of this nature could instigate related research for dissertations.

Sharing of this type of teaching experience by articles in professional journals or discussions at appropriate national meetings would be of real value for interested workers in the field. The sharing of such experiences should be of value in furthering our joint efforts in such a direction.

The one research area I would like to focus on in the remaining section of the paper deals with the nuclear family's impact on the personality development of the child.

As the primal social environment of the developing child, the intensive study of nuclear family interaction and interrelationships holds the key to much of our understanding of the developing child, particularly in the earliest years of growth. Although much research has been done, we need to know much more than we do about the complex involvements, interactions and relationships of family members upon each other through the life cycle. In so doing we can ascertain the factors influential in affecting the course of personality development.

The family is viewed as one complex environmental variable having the highest frequency of points of contact, at least in the early years of the child's life. Some major questions then arise in our attempts to describe these contacts, to indicate their effect on the child at varying points in his development. In Chein's (3) terms, for example, analysis of the family, in its structure-function organization as an environmental variable supporting, constraining or directing the child's growth, should yield valuable data as to how and in what ways the family influences the child's developing personality.

The steps in this programmatic research suggestion are similar to what was mentioned previously about the whole field.

1. What is needed is increased descriptive data on the functions of various family members as they interact at various stages of the individual's development. In previous research in parent-child relationships, we have dwelt too heavily on the mother-child relationship. Important as this relationship is for the child, let us not ignore the male member of the family. Ironically, in our Freudian psychology the father plays a prominent role, yet in our actual research in parent-child relationships we have tended to ignore the role of the father or to study his role reconstructively, using mothers and children as informants. Thus, research should be concerned with the

father's role in his relationship with his child in its many aspects. Such problems then emerge, as the type and quality of contacts the father has with his children, the differential effects of such experience for the father and how these change with changes in the child. The mother's role should of course be examined in a comparable fashion. Much, of course, has been done as far as the mother is concerned. Both parents should be necessarily involved.

The next step here is to study comparatively and interactively fathers and mothers in terms of relative effect and influence. Thus, we are interested in the relative influence of parents in control of the family and of the children. What is the effect on the developing child of inconsistency between parents? Further, what difference does it make who is the frustrator, the gratifier, the indifferent one in relation to the child's need satisfactions? In other words, we have to focus on the parents comparatively to assess and evaluate the type and kind of points of contact they have with the child.

2. There is also great need to study the other members of the nuclear family, especially siblings. The interactions that occur within the family that have potential significance for the child stem from siblings as well as from parents. The siblings in their relationship to each other may reinforce, positively or negatively, parental actions, and consequently may function to maintain or to change potential responses of a child to his parents. For example, when an older sibling functions to gratify the love needs of a younger sibling, while the parent is frustrating such gratification, would we not expect different personality characteristics to appear as compared to the situation where no sibling performs the love-giving function?

The analysis of the family as it functions in various ways to effect the child should follow a longitudinal pattern in order to assess its role at different developmental stages of the child. As the child matures and his familial relationships begin to function differently, the influence of the nuclear family may vary. Thus, we have to view the two variables here, the child and the family, as two changing systems whose interactive and influence patterns will shift in relation to different demands and needs of each. As we follow this trend, seeking specific functional relationships, we should be able to specify in necessary detail those variables most readily influenced by familial experiences.

Families in such research should be drawn so as to be more representative of the population as a whole. We need to devise ways to contact and to maintain rapport for research purposes with lower income families, with families of different ethnic groups, and in general, nonclinical populations of families, in order to get the variations of family functions and their different influences on the child.

The more intensive our study of the family, and the sets of conditions residing therein, the greater should be our knowledge of the individual differences of developing children.

The relationships that exist within the family are to a large measure expressed in the type of interactive behaviors that occur among family members. Feelings of love, respect, trust, or the negative aspects of these take some concrete behavioral form. Then it behooves us, as one point of emphasis, to study the effect of these overt behaviors. What family members say and do to each other in the course of their living together should be evaluated in terms of impact on the family participants.

Within this context of behavioral investigation I should like to emphasize the urgent need for intensive study of the role of parent behaviors in influencing the numerous aspects of the child's development. Whereas we at Merrill-Palmer are focusing on parental control behavior, there are many other categories of parent behavior that are relevant, e.g., play, social interaction, etc.

Thus, we select out one aspect of the family environment, the behavioral one, and must set out to describe parent behavior in relation to various experiential conditions, e.g., what does the parent do when the child comes for comforting, or what does the parent do when the child is frustrated? But we must also be concerned with how the parent changes in behavioral interaction with the child as the child gets older, when the issues and needs about which parent-child interactions occur may take more complex forms.

In our study of parental behavior, we have many complex theoretical and methodological problems. A major one seems to me to revolve about the effective stimulus variables within the behavioral act. What is it that stimulates the child to respond? For example, what is a joking statement and said in a joking tone with joking words may be in reality a threatening anxiety-producing stimulus to the child. Is it the fact of joking, is it the timing of the statement, is it the way the parent looks, or what?

In examining such questions as to the effect of the stimulus, it seems that we must include as intervening variables the child's perceptual and cognitive functions. Thus, the study of the behavior of parents toward their children points to the development of a theory in which such systems are interrelated. This includes the acts of parents, the responsive process in the child, and the structure of the stimulus itself. We will have to integrate concepts from various psychological subsystems, e.g., personality, cognition, etc., in order to explain adequately the role such behaviors play at different developmental stages.

In addition to focusing on the behavior of the parents, we must not ignore the influence of the child's behavior on the parent's behavior. In other words, this interaction is a mutual one and hence some of the contemporaneous antecedents of each set of behaviors lie within the behaviors of the other person. Thus, the child's acts stimulate the parent and the parent's acts stimulate the child. When we look toward the role child behaviors have in influencing the parents, we again need to ascertain the meaning these child behaviors have for the parent and what properties of child behavior instigate various adult responses. Here we become involved in the value system, the motivational system and the cognitive system of the parent, as antecedents for what child behavior will be response producing and for the nature of the parental response.

Up to this point we have examined the family from within and how its members influence and affect each other. But we can and should study the family as a unit and its relationship to its larger environment. Since our society is an ever-changing thing, since it is true that the growing child in his family is dynamically changing, then we are faced with a research problem involving the relationships of two dynamically changing systems. Here we are interested in the effect of these two variables on each other. What is the effect of a changing society with shifting values, with complex requirements for achievement, and with undefined customs and mores, on the developing child and his family? How do these societal forces influence the child's adjustment, what and how many adjustments does a child have to make, what happens to the child and the family in their sets of expectations of behavior from others, to achievement in life, to peer and other social relationships? Finally, what is the effect of these societal instabilities on the relationship between children and their parents?

Within this complex society, we have to examine the influence on the child and his family of other individuals, other relationships. One microcosm that is of importance in studying the factors involved is the neighborhood. This is particularly true in urban areas, where this geographical unit is to a large measure the habitat of the young child. Once the child starts to move in it, people other than family members are going to play a part in influencing him directly or indirectly. His experiences in these situations may bring new problems into the home which were never anticipated by parents. Although these may seem at this point spurious and unpredictable forces impinging on the child and his family, the investigation of the world of the child to include these variables would aid in throwing some light on the role of ecological forces in the development of the child.

I would prefer this type of approach at this time. We frequently talk of cultural forces when we speak of the influence of the social environment on the child. This is a very gross point of view, particulaly in a complex and heterogeneous society as ours. We have to ask the question, what unit of culture are we to study? In our society, culture and social class are extensive abstractions and necessary concepts in the broad social picture of examining human growth and development. Neighborhood is a utilitarian concept. No doubt it reflects many of the broad cultural patterns, but at the same time it selects and shapes these into its own image and it is this unique microcosm to which the child is making some of his early adaptations. For example, materialistic values may be said to be dominant in this country, but it is the interpretation of these in a particular neighborhood and the social consequences of the locally approved pattern that is having its impact on the children under study. Thus, by focusing on field studies which specifically deal with home-neighborhood-school constellations we gather the specifics which are necessary in order to make our adequate generalizations. An excellent example of the kind of approach I am advocating here is the work of Barker and Wright (2) in their studies of the ecological habitat.

These questions and problems, although briefly stated are obviously extensive in scope. Yet I feel a many-pronged attack on various aspects of them is of vital importance in our understanding of the development of the child within his psychosocial environment.

The practical and theoretical contributions made by such study are numerous. For as we improve our understanding of the intricacies of the developmental process within the family and its social environment, it should help us to understand the patterns that exist, the deviations that occur and eventually the therapeutics needed to maintain a healthy course of change and growth of the child and his family.

REFERENCES

1. BALDWIN, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.
2. BARKER, R. G., & WRIGHT, H. F. *Midwest and its children.* Evanston, Ill.: Row, Peterson, 1955.
3. CHEIN, I. The environment as a determinant of behavior. *J. soc. Psychol.,* 1954, 39, 115-127.
4. GESELL, A., et al. *The first five years of life: A guide to the study of the pre-school child.* New York: Harper, 1940.
5. PIAGET, J. *Language and thought.* New York: Humanities Press, 1951.
6. SHIRLEY, M. M. *The first two years: a study of twenty-five babies. Vol. I. Postural and locomotor development.* Minneapolis: Univer. of Minnesota Press, 1931.

EXPERIMENTS WITH CHILDREN ON THE HYPOTHESES OF ACQUIRED DISTINCTIVENESS AND EQUIVALENCE OF CUES

Charles C. Spiker
Iowa Child Welfare Research Station

The participants in this conference have been asked to concern themselves with matters of high priority with respect to research in the field of child development. This writer takes the position that research problems lending themselves to integration within existing theory, or which contribute directly toward the construction of theory, are among the most fruitful problems that can be investigated at any given time. Since the writer views the field of child development as a multi-disciplinary field, since he is convinced that few scientists are genuinely competent in more than one or possibly two disciplines, and since his own training has been in the psychological area, he prefers to restrict the range of his comments to the area of child psychology.

It does not seem that the child can profitably be viewed as a unique subject for psychological investigation. Rather, the child is another subject for psychological study, and, for theoretical and/or practical reasons, certain problems are more appropriately studied with children than with other subjects. The results of such research with children are thus expected to contribute to a general theory of behavior and not to theories of child behavior. This conception does not deny that productive research workers with children must have some formalized training in the handling of children. But it would argue that however important, socially and morally, we may consider such training, it does not differ fundamentally from the type of training that is provided the student of rat behavior. In particular, the necessity for such training does not vitiate the equally important need

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

for training in the areas of the philosophy of science, research methodology, experimental design, statistical analysis, and general psychology.

The remainder of this paper is largely concerned with an illustration of the manner in which psychological research with children can contribute to general behavioral theory. It will present a review of some recent research with children dealing with two specific hypotheses from behavior theory.

The first of these has been variously referred to as the hypothesis of secondary generalization (14), mediated generalization (3), or the acquired equivalence of cues (18). In brief, it states that if a subject has been trained to make the same (mediating) response to two or more dissimilar stimuli, there will be an increased tendency to generalize to the other stimuli other (mediated) responses *subsequently* learned to one of the stimuli. This hypothesis is derived from familiar conditioning principles plus some specific assumptions concerning the occurrence and function of response-produced stimulation. The derivation may be sketched as follows. Learning the mediating response (R_1) to each of several dissimilar stimuli ($S_1, S_2, \ldots S_n$) assures that on subsequent presentation of these stimuli, the stimulation (s_1) produced by R_1 will be present. Any other response (R_o) subsequently learned to one of these stimuli (S_1) will also be conditioned to s_1. Thus, if any other of the stimuli ($S_2 \ldots S_n$) is presented, R_1 will occur, together with s_1, and s_1 will tend to evoke R_o.

The second hypothesis has been referred to as the principle of acquired distinctiveness of cues (18). As the name implies, this hypothesis has to do with discrimination rather than with generalization, although the basic assumptions involved in the two hypotheses are highly similar. Briefly, the hypothesis is that if a subject learns to respond differentially to two or more similar stimuli, he will subsequently learn more rapidly other discriminal responses to these stimuli. The learning of the first set of discriminal responses ($R_1, R_2, \ldots R_n$) to the stimuli ($S_1, S_2, \ldots S_n$) assures that on subsequent presentations of these stimuli, the stimulation ($s_1, s_2, \ldots s_n$) produced by the discriminal responses will be present. Hence, when S_1 is presented, the total stimulus complex will include s_1; when S_2 is presented, the total complex will include s_2; and so on. Assuming that s_1 is more discriminable from s_2 than S_1 is from S_2, the learning of the first set of discriminal responses should serve to reduce the functional similarity among the stimulus complexes. One of the implications of this set of assumptions is that a new set of discriminal responses ($R_a, R_b, \ldots R_z$) could be learned more readily to the stimulus set ($S_1, S_2, \ldots S_n$) if the subjects have first learned other discriminal responses than if they have not. This implication follows from the familiar conditioning law that relates the difficulty of a discrimination learning task to the similarity of the discriminal stimuli.

Some of the more convincing empirical evidence for these hypotheses has been obtained with children. One of the earliest investigations of the hypothesis of acquired equivalence of cues (AEC) was conducted by Birge (3). Her grammar school subjects first learned to call two of four nonsense animals by one name and the other two by another name. One stimulus from each name-pair were presented together and the subjects learned to find a piece of candy under one of them. Finally, the other member of each name-pair were presented together and the subjects were allowed to choose one of them. She found that the majority of subjects chose that stimulus which had the same name as the stimulus under which they had previously found candy. Moreover, the tendency to choose in this direction was greater for subjects who were required to verbalize the names overtly during the candy-finding and test stages than for subjects not required to verbalize overtly during these stages. In a further experiment, Birge found out that subjects were more likely to generalize a name from one stimulus to another if they had found candy associated with both stimuli than if candy had been associated with only one of them. Shepard (24) replicated a part of Birge's first experiment, using different stimuli, and found essentially the same results.

An experiment by Jeffrey (15) with preschool children provides another demonstration of the role of mediating responses in generalization. He trained his subjects to move a lever in one direction to a white stimulus and in the opposite direction to a black stimulus. He then taught some subjects to call a gray stimulus "white" and others to call it "black," and next retrained the subjects on the lever moving task to the white and black stimuli. Following this, he presented the gray stimulus interspersed with black and white stimului to determine how the subjects would respond to the gray stimulus. He found that if the subjects had been taught to call the gray stimulus "white," they responded to it as they did to the white stimulus; if they had been taught to call the gray stimulus "black," they responded to it as they did to the black stimulus. He found a similar tendency, though not so strong, for another group of subjects who instead of using the names "black" and "white," had been taught to turn the handle to the right or to the left for the white, black, and gray stimuli. Thus, like Birge's experiments, Jeffrey's indicates that either verbal or motor (manipulative) responses may serve as mediating responses.

The generality of the AEC hypothesis is suggested in a recent experiment by Eisman (9) in which she produced color preferences in children through mediated generalization. She first taught grammar school children a different name for each of three white geometric blocks. She then reinforced them for selecting a particular one of the three blocks. Next, she taught the children the same three names for three differently colored blocks. When the subjects were then asked to find the reward under one of the colored blocks, they tended to choose that color which had been assigned the same name as the rewarded geometric block. By giving additional naming experience with colored jar tops and glasses, she was able to demonstrate a selective preference for that one of three white blocks that had been assigned the preferred color name, and also a verbalized preference for that one of three fictitious groups of children that had been assigned the preferred color name.

Results consistent with the AEC hypothesis have also been reported with adult subjects by Murdock (19), Bugelski and Scharlock (4), and Russell and Storms (23). In the latter two studies, it was reported that the subjects were unable to verbalize the basis of mediation.

The empirical evidence for the hypothesis of acquired distinctiveness of cues (ADC) is somewhat more difficult to interpret. The simplest paradigm involves two groups of subjects given a pretraining and a transfer task. The two groups are given the same transfer task under identical conditions, but differ with respect to the type of pretraining. The experimental (relevant stimulus) group is given preliminary training in which distinctive responses (usually names) are learned for the stimuli involved in the transfer task. The control (irrelevant stimulus) group is given similar training with stimuli that are not involved in the transfer task. The transfer task itself requires both groups to make other discriminal responses to the stimuli. This paradigm is assumed to control for positive transfer attributed to such non-specific factors as "warm-up" or "learning to learn." Thus, the relevant stimulus group is expected to learn the transfer task more rapidly than the irrelevant stimulus group due to the contribution of the response-produced stimulation to the distinctiveness of the total stimulus complexes.

The results of research with adults on the hypothesis have been conflicting. Arnoult (2), J. Cantor (6), Gerjuoy and Bechtoldt (12), McAllister (17), and Robinson (21) were unable to show facilitation as a function of verbal pretraining on the discriminal stimuli. J. Cantor (7), Dysinger (8), Gagné and Baker (10), Goss (13), Price and Lewis (20), and Rossman and Goss (22) did demonstrate facilitation rather conclusively. Examination of the experimental conditions that obtained in these experiments, together with analysis of the theoretical assumptions underlying the ADC hypothesis, suggests at least four conditions which should be met in a test of the hypothesis.

First, the discriminal stimuli must be sufficiently similar that generalization initially occurs from one to the other; at the same time, they must be sufficiently dissimilar that discriminal responses can be learned to them. *Second*, the discriminal stimuli must be of such a nature as not to elicit verbal names or descriptive phrases among the control subjects. If this condition is not met, the experimental subjects will have no advantage with respect to the experimentally introduced names, and may even be handicapped by competing names. It is this condition that seems to make the child the ideal subject for testing the ADC hypothesis, since he is less likely to use descriptive phrases than is an adult and the selection of appropriate stimuli is thus made easier. *Third,* the responses to be learned in the transfer task should be simple and discrete. If these responses involve mastery of a complex motor coordination, the facilitation contributed by the pretraining may be masked or at least contribute only slightly to the total variation in the learning scores. *Fourth,* if the positive transfer is to be attributed to the effects of the names learned, there should be an adequate control for the "warm-up" or "learning to learn" effects.

A recent experiment to test the ADC hypothesis was conducted by Gerjuoy (11) with elementary school children. Her transfer task consisted of the successive presentations of three stimulus lights, differing from each other only in position on a 20 in. by 20 in. panel, which the child learned to turn off by pushing the appropriate one of three response buttons arranged in a horizontal row. Several experimental groups were given a preliminary task in which they were required to learn nonsense names for the three light positions used in the transfer task. Reasoning that the more similar the names in pretraining, the less effective they should be in the transfer task, Gerjuoy required one experimental group to learn highly similar names while another learned dissimilar names. Some subjects in each of these two groups were required to verbalize the names overtly in the transfer task prior to responding while others were not required to do so. Control subjects were given a preliminary task in which they were simply required to call off the names of colored lights as they appeared on the stimulus panel. Gerjuoy found that all subjects given the name-learning as the preliminary task performed better on the transfer task than those given the color-naming task, but there were no differences in this respect among the experimental groups.

An experiment with preschool children by G. N. Cantor (5) is also relevant to this hypothesis. His transfer task was a simultaneous discrimination problem that involved reinforcement for the selection of one of two toy cars differing from each other with respect to the stimuli pasted on one side. The stimuli were sketches of the faces of two girls with marked resemblance. Relevant stimulus subjects were given pretraining in which they learned to call one of the faces by one name and the other by another name. A second ("attention") group did not learn names for the faces, but had an equal number of trials in which they pointed to various parts of the faces (nose, eyes, etc.), thereby assuring "perceptual" experience with the stimuli. An irrelevant stimulus group learned names for a pair of boys' faces. All three groups were immediately given the transfer task. Cantor found that the group given pretraining with the relevant stimuli learned the transfer task significantly faster than either of the other two groups, and that these latter two groups did not differ significantly from each other.

A subsequent experiment by Norcross and the writer, using Cantor's stimuli, was conducted to determine whether discrimination experience with the relevant stimuli is sufficient to account for the facilitation. The experimental design was like Cantor's except that in place of his "attention" condition, a condition was used under which the subjects were required to learn to say "same" when two identical pictures were presented and to say "different" when the two pictures were different. This procedure assures that the subjects must discriminate between the two stimuli without necessarily giving different names to them. The relevant group learned significantly faster than the irrelevant and discrimination groups, and the latter two did not differ from each other.

This series of experiments, with those conducted on adult subjects, confirms the usefulness of the ADC hypothesis. The experiments by Gerjuoy (11) and others (7, 8, 13, 22) indicate that a facilitation occurs that cannot be attributed to warm-up or a learning set induced by the preliminary task. The experiment by G. N. Cantor (5) indicates that this facilitation cannot be attributed to preliminary visual or perceptual experience with the stimuli, and that by Norcross and the writer suggests further that preliminary discriminatory experience is insufficient to account for the subsequent facilitation. The most important contribution of these experiments lies in the control for perceptual and discriminatory experience. Probably only with young children could one give discrimination experience without at the same time eliciting consistent names or verbal descriptions for the stimuli.

The facilitative effects of stimulus pretraining with names has been shown in the delayed reaction experiment (28). The experimental group was given the usual relevant stimulus pretraining in learning distinctive names for each of a pair of highly similar sets of concentric circles. The control group was given pretraining with these stimuli that involved saying "same" when the two stimuli presented were identical and saying "different" when they were different. The transfer task required the subject to select which of the two stimuli he had seen baited with a reward 25 seconds before. The stimuli were not visible to the subject during the delay period. It was found that subjects given the verbal pretraining selected the correct stimulus significantly more frequently than those without such training, and that this effect was significantly greater for younger than for older subjects. These results are consistent with the ADC hypothesis, as well as with the hypothesis that the names serve as substitutes for the absent stimuli during the delay period.

Assuming tentative verification of the ADC and AEC hypotheses, Shepard (25) made an interesting application of these hypotheses to the classificatory behavior of preschool children. Her transfer task involved five nonsense figures, and the subjects were rewarded for pulling a lever to three of them and were not rewarded if they pulled the lever to the other two. The relevant stimulus group was given a pretraining task in which they learned one name for the first three stimuli and a different name for the other two. An irrelevant stimulus group was given similar training, but with stimuli not used in the transfer task. The relevant stimulus group made fewer errors on the transfer task than did the irrelevant stimulus group. Shepard accounted for this difference in terms of the ADC and AEC hypotheses. Learning the same name for the stimuli within a class is assumed to increase the amount of generalization that occurred among them. Learning different names for the members of the two different classes is assumed to decrease the amount of generalization between members of the different classes. Both processes thus contribute to facilitation of the relevant stimulus group.

Shepard's application of these two principles to classificatory behavior suggests an explanation for another well-established phenomenon. Kuenne (16) used preschool children to study the transposition phenomenon. As a point of departure, she used Spence's discrimination learning theory (26, 27), which predicts transposition for animals in the two stimulus problem for test stimuli that are similar to the training stimuli, but little or no transposition for test stimuli that are distant from the training stimuli. Kuenne suggested that subjects who are articulate with respect to the stimulus differences may use these verbal responses in such a way that transposition to more distant test stimuli also occurs. She was able to demonstrate, under specified conditions, that the distance of the test stimuli from the training stimuli may be greater, and transposition still occur, for children with higher mental ages (and therefore presumably more articulate with reference to the stimulus relationship) than for children with lower mental ages. A similar finding was reported by Alberts and Ehrenfreund (1) in which the range of transposition was found to be an increasing function of chronlogical age.

Spiker, Gerjuoy, and Shepard (29) conducted a related experiment in which transposition was studied in the intermediate size problem. They found the amount of transposition to be related to the subject's knowledge and understanding of the concept "middlesized." That is, subjects who knew and understood the concept "middlesized," as determined by a preliminary test, transposed to a considerable degree, but subjects who did not understand the concept failed to transpose to any great degree.

The ADC and AEC hypotheses may be used to analyze the manner in which the verbal concepts affect transposition behavior. Consider the two stimulus problem in which the subject is reinforced for selecting the larger of two stimuli. Suppose that the larger stimulus elicits, by virtue of previous learning, the verbal response "larger one" (or its linguistic equivalent) and the smaller stimulus elicits "smaller one." Suppose, moreover, that for this subject, the larger and smaller stimuli of any pair of stimuli with discriminable differences in size elicit these distinctive verbal responses. According to the AEC hypothesis, any response (e.g., a choice response) learned specifically to the larger stimulus of a given pair should therefore generalize to the larger stimulus of any other pair, regardless of the distance of the test pair from the training pair. And, because the two stimuli elicit distinctive responses, the generalization between them should also be decreased. Thus, extensive transposition of articulate subjects would be explained by mediated generalization and acquired distinctiveness of cues. A similar derivation can be made for the intermediate size problem.

This account does not, of course, explain how verbal names are learned for stimulus relationships in the first place. Considerable research with young children on such problems seems advisable before speculation is likely to be profitable.

The results of the experiments reported above indicate that research with children can be brought to bear on hypotheses from general behavior theory. Obviously, children are more appropriate as subjects for some problems than for others. There are many factors to consider in selecting problems for research with children. For example, there are some phenomena that are more or less unique, or at least occur in their purest form, in children. The development of the first linguistic behavior can be studied only in very young children. Imitation learning can be studied in children without many of the complicated processes that confound adult imitation. The use of children provides a means of selecting subjects with a level of language ability more under the control of the experimenter than would be true for adults. To arrange for adults to show the same gradual learning as children show in simple problems, the task must often be made many times more difficult at the risk of masking the basic processes. The assumption that the subject is unaware of the purpose of an experiment, the tenability of which is central in numerous experiments, can often be more plausibly maintained if the subjects are unsophisticated children. The development of experimental tasks, for which generalization from previous experiences will be minimal, is undoubtedly easier to accomplish for child than for adult subjects. Certain pragmatic considerations also suggest the more effective use of children in much psychological research. If one is interested in the effects of incentive on performance, dime-store toys for children may be as good or better than dollars for adults. Or, the observable effects of a mild experimental frustration on child behavior may be greater than those of an elaborately arranged "traumatic" experience on adult behavior. Or again, children can usually be more readily convinced than adults that their performance represents success or failure.

Such considerations may result in economy for the experimenter with the ingenuity to make use of them. It may be these factors that many have in mind when they insist that the child mirrors the basic psychological processes more faithfully than does the adult. Certainly these considerations suggest that, given carefully developed apparatus, procedure, and research methodology, the child may be one of the most fruitful subjects for psychological experimentation, not only in the investigation of practical problems as in the past, but also in the testing of hypotheses that stem from general psychological theory.

REFERENCES

1. ALBERTS, E., & EHRENFREUND, D. Transposition in children as a function of age. *J. exp. Psychol.,* 1951, 41, 30-38.
2. ARNOULT, M. D. Transfer of predifferentiation training in simple and multiple shape discrimination. *J. exp. Psychol.,* 1953, 45, 401-409.
3. BIRGE, J. S. The role of verbal responses in transfer. Unpublished doctor's dissertation, Yale Univer., 1941.
4. BUGELSKI, B. R., & SCHARLOCK, D. P. An experimental demonstration of unconscious mediated association. *J. exp. Psychol.,* 1952, 44, 334-338.
5. CANTOR, G. N. The effects of three types of pretraining on discrimination learning in preschool children. *J. exp. Psychol.,* 1955, 49, 339-342.
6. CANTOR, JOAN. The effects of pretraining on the performance of a simplified Mashburn task. Unpublished master's thesis, State Univer. of Iowa, 1953.
7. CANTOR, JOAN. Amount of pretraining as a factor in stimulus predifferentiation. *J. exp. Psychol.,* 1955, 50, 180-184.
8. DYSINGER, D. An investigation of stimulus predifferentiation in a choice discrimination problem. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.
9. EISMAN, BERNICE S. Attitude formation: The development of a color-preference response through mediated generalization. *J. abnorm. soc. Psychol.,* 1955, 50, 321-326.
10. GAGNÉ, R. M., & BAKER, K. E. Stimulus predifferentiation as a factor in the transfer of training. *J. exp. Psychol.,* 1950, 40, 439-451.
11. GERJUOY, I. R. Discrimination learning as a function of the similarity of the stimulus names. Unpublished doctor's dissertation, State Univer. of Iowa, 1953.

12. Gerjuoy, H., & Bechtoldt, H. P. Generalization effects in successive discrimination learning as measured by response similarity. Department of Psychology, State Univer. of Iowa, 1955.

13. Goss, A. E. Transfer as a function of the type and amount of preliminary experience with task stimuli. *J. exp. Psychol.,* 1953, 46, 419-428.

14. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.

15. Jeffrey, W. E. The effects of verbal and non-verbal responses in mediating an instrumental act. *J. exp. Psychol.,* 1953, 45, 327-333.

16. Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.,* 1946, 36, 471-490.

17. McAllister, D. E. The effects of various kinds of relevant verbal pretraining on subsequent motor performance. *J. exp. Psychol.,* 1953, 46, 329-336.

18. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.,* 1948, 43, 155-178.

19. Murdock, B. B., Jr. Effects of failure and retroactive inhibition on mediated generalization. *J. exp. Psychol.,* 1952, 44, 156-164.

20. Price, H. G., & Lewis, D. Increased pronouncing behavior as a factor in serial learning. *J. exp. Psychol.,* 1954, 47, 95-100.

21. Robinson, J. S. The effect of learning verbal labels for stimuli on their later discrimination. *J. exp. Psychol.,* 1955, 49, 112-114.

22. Rossman, I. L., & Goss, A. E. The acquired distinctiveness of cues: The role of discriminative verbal responses in facilitating the acquisition of discriminative motor responses. *J. exp. Psychol.,* 1951, 42, 173-182.

23. Russell, W. A., & Storms, L. H. Implicit verbal chaining in paired-associates learning. *J. exp. Psychol.,* 1955, 49, 287-293.

24. Shepard, W. O. Mediated generalization with high interstimulus similarity. Unpublished master's thesis, State Univer. of Iowa, 1953.

25. Shepard, W. O. The effects of verbal pretraining on discrimination learning in preschool children. Unpublished doctor's dissertation, State Univer. of Iowa, 1954.

26. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, 43, 427-449.

27. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.

28. Spiker, C. C. Stimulus pretraining and subsequent performance in the delayed reaction experiment. *J. exp. Psychol.*, in press.

29. Spiker, C. C., Gerjuoy, I. R., & Shepard, W. O. Transposition by children with the intermediate size problem. *J. comp. physiol. Psychol.*, in press.

PSYCHOLOGICAL DEVELOPMENT IN MIDWEST

HERBERT F. WRIGHT
University of Kansas

Midwest is a small town in a midwestern farm region of the United States. Some 100 of its approximately 700 inhabitants are children under 12 years of age. Although the supply lines of American culture and technology run through it, Midwest nonetheless has essential independence and completeness (2, pp. 20-32), with the result that it holds the lives of these children almost entirely within its borders. One confronts an "intelligible field of study" (4), therefore, in asking the question raised here: What age related changes occur in the behavior and psychological living conditions of Midwest children? What is it like to grow up in Midwest?

The Midwest Field Station has been working for some time with this and related problems in psychological ecology.[1] Its special task has been to describe along as many dimensions as can be accommodated by available procedures and concepts the psychological habitat and naturally occurring behavior of Midwest children. This it has aimed to do with something of the outlook and aspirations of the weather outpost and of biological and astronomical field stations which have reaped from continuing records of naturally occurring phenomena substantial benefits in positive knowledge and basic theory. Methods of the research and results of studies to date have been reported elsewhere (see especially 2, 3, 11, 12). I will try here only to review some of these methods and to bring together some of their

---

[1] The work reported here has been directed by Roger G. Barker and the author, and supported in part by the Institute of Mental Health of the National Institutes of Health, U.S. Public Health Service, and the University of Kansas. Other direct contributors to the studies represented include William A. Koppe, Jack Nall, and Phil Schoggen.

CHILD DEVELOPMENT, Vol. 27, No. 2 (June, 1956)

results that answer only to the one problem of change with age in behavior and its conditions in Midwest.

DEVELOPMENTAL CHANGE AND BEHAVIOR SETTINGS

Behavior Settings as Community Units

Procedures to be reviewed below indicate that of their person hours in one recent year Midwest citizens of all ages spent

18,525	in Clifford's Drug Store,
7,600	in the Second Grade Classroom,
7,300	in Hooker's Tavern,
3,750	in the Methodist Regular Worship Service,
1,000	in the Midwest Cemetery,
540	in the City Library, and
344	in the Brownie's Regular Meeting

of this community. Here are plainly visible Midwest centers of human activity. We have called such action areas *behavior settings,* and we have used them to describe some aspects of psychological development in Midwest.

A behavior setting can be defined as a stable part of the physical and social milieu of a community together with an attached standing pattern of human behavior. Always the milieu part has its own distinctive characteristics of time, place, things, and personnel that accommodate the behavior pattern. Always, too, this pattern is extraindividual and invariant in that it stays essentially the same while different persons come and go.

Take as a single example the Methodist Regular Worship Service. Its time is 11 to 12 a.m., every Sunday, and its place the Methodist Sanctuary. Its things include hymn books, collection plates, and a pulpit. Ushers, congregation members, the choir, and a minister are it personnel. These elements form the milieu part. The behavior pattern includes ushering, sitting, standing for prayer, preaching and listening, gathering in the vestibule, conversing, and shaking hands, all of which change little Sunday after Sunday, year after year, no matter who goes to church.

Behavior settings are common empirical facts of such prominence that independent observers identify them and most of their features and functions with near unanimity. Yet each contains forces that coerce individual behavior, as abrupt changes in the action streams of individuals when they leave one Midwest setting and enter another regularly imply. Collectively, moreover, a community's settings have a kind of omnipotence in that virtually all things done by its citizens occur within their limits. For these and other such reasons (3, pp. 45-57) we have found in behavior settings appropriate analytical units for research on some aspects of the natural history of human behavior in Midwest.

A Behavior Setting Survey

All of Midwest's behavior settings, from New Year's Holiday Dinner to Santa Claus at the Court House, occur at least once in the course of a year. The station has accordingly made a behavior setting survey of Midwest for a town year: July 1, 1951, to June 30, 1952. There were identified during this survey year 2,030 Midwest settings, of which 1,445 occurred in homes and 585 in more public areas. Special criteria were used to verify the existence of different settings as discrete units of the same magnitude via tests of their mutual independence (3, pp. 62-66).

Chief consideration will be given here to the 585 public units or *community behavior settings* in which, according to our data, Midwest children spent approximately one-third of their waking time. These were grouped into 107 *varieties*, among which there were *School Classes, Indoor Entertainments, Restaurants and Taverns, Barber and Beautician Shops, Sunday School Class Meetings, Circuses and Carnivals, Organizational Meetings With Refreshments, Organizational Meetings Without Refreshments, Parades, Funeral Services, Food and Rummage Sales, School Field Trips*, and *Fire Drills*. Seventy of the 107 varieties and approximately 205 member settings were found to occur in the town during an average week day.

TABLE 1

AGE GROUPS OF THE MIDWEST POPULATION

	Number of Persons		
	Males	*Females*	*Total*
INF: Infant (birth to 1 year, 11 months)	15	10	25
PRE: Preschool (2 years to 5 years, 11 months)	18	12	30
YNG: Younger School (6 years to 8 years, 11 months)	14	15	29
OLD: Older School (9 years to 11 years, 11 months)	14	21	35
ADO: Adolescent (12 years to 17 years, 11 months)	25	25	50
ADU: Adults (18 years to 64 years, 11 months)	172	204	376
AGD: Aged (65 years and over)	82	80	162
All ages	340	367	707

NOTE.—The different age groups are identified in subsequent figures by the abbreviations used here.

The procedures of the behavior setting survey can be used to describe aspects of the behavior and life situations of adolescents and adults as well as of children. Therefore, to get information on behavioral change throughout the life span in Midwest, the station has applied these procedures to all of the seven age groups of Table 1, with results now to be sampled.

Time in Settings and Age

We can begin with a plain question. What changes occurred with age in the time Midwesterners spent in different behavior settings?

The answering data are based on estimates of *occupancy time,* anticipated by the itemization above of person hours in selected settings of Midwest. This measure was obtained by multiplying the number of occurrences of a setting during the year (its frequency, f) by the number of Midwest citizens who entered it during each occurrence (its population, p) by the time in hours these persons spent in the setting when they entered it (its duration, d). The product fpd gave the person-hours-per-year or the occupancy time of the setting. Separate occupancy measures were computed for each age group and other subgroups of the population. Also, the sum of the time spent in all community settings was then determined for every such group. Finally, the percentage contributed to this sum by each setting was computed as a group *occupancy index* for that settings unit.

TABLE 2

OCCUPANCY INDEXES OF THE TEN COMMUNITY BEHAVIOR SETTINGS IN WHICH MIDWEST PRESCHOOL CHILDREN AND MIDWEST ADULTS SPENT THE MOST TIME

	OCCUPANCY INDEX: PER CENT OF		
	22,420	593,994	
Trafficways	21.4	8.1	Trafficways
Culver Shoe Repair	6.7	2.9	Telephone Office
Midwest Theater	6.2	2.6	Garnett's Grocery
Mrs. H Baby Care	4.5	2.6	Kanes Grocery
Methodist Kindergarten	3.3	2.5	Gwynn Cafe
Mrs. D Baby Care	3.3	2.5	Bank of Midwest
Presbyterian Kindergarten	2.0	2.4	Wherry Door Company
Clifford's Drug Store	1.7	2.4	Midwest Theater
Pearl Cafe	1.6	2.4	Clifford's Drug Store
Saturday Night Dance	1.3	2.1	Midwest Hardware
Preschool Children			*Adults*

NOTE.—Trafficways are streets and sidewalks. The Culver Shoe Repair housed the Culver family whose eleven (11) children had many friends. The kindergartens are Sunday-School classes. Children often tag along to dances in Midwest.

Table 2 lists, first, the 10 community settings in which Midwest preschool children spent the most time and, second, the 10 in which Midwest adults spent the most time. Each setting is named with its occupancy index. Some impression as to what growing up brings in Midwest may be gained by comparison of the two lists in this table.

Our occupancy figures indicate that Midwesterners spent 5,130,000 hours (waking and sleeping) in family settings and 1,030,658 hours in community settings. Data on the average number of hours per day all residents occupied these Midwest living areas are summarized for seven age levels in Figure 1. Each average is for all weekdays, holidays, and Sundays of the entire year.

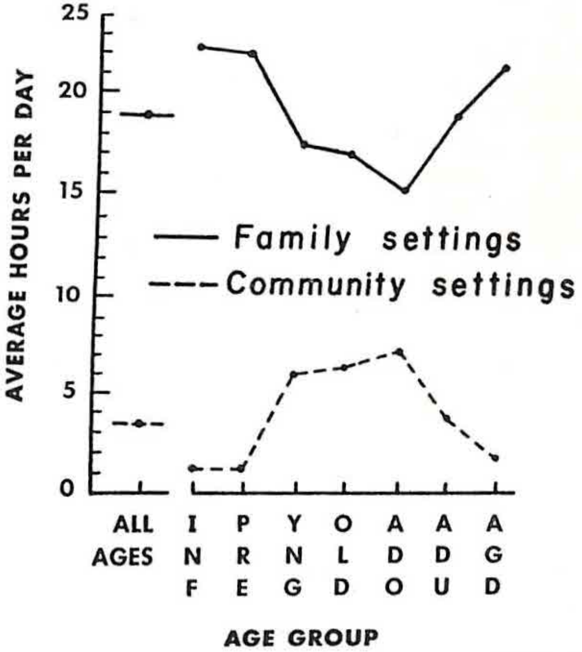


FIGURE 1—Average hours per day spent in family and community behavior settings by the total population of Midwest and by different age groups.

Figure 1 shows that family settings surpassed community settings about 5 to 1 in competition for the time of Midwest citizens. But it also shows dramatic change with age in this competition. Temporal preeminence of the family declined greatly from infancy and early childhood through adolescence, but then rose above the maximum level for school children in adulthood, and approached the level for infants in old age. One might easily be surprised at first to find older children and adolescents spending, as Figure 1 indicates, much more time than their elders away from home. Are not the years of later childhood and youth family years? The answer, of course, is that they are school years as well. It is clear in any case that growing up in Midwest through adolescence was considerably a matter of growing into public behavior areas of the community.

Number of Settings and Age

What changes occurred, then, as Midwesterners grew older in the number of different community settings they entered? How did freedom to roam over the town vary with age?

This question the station has answered for each of several Midwest population subgroups in two steps. The *territorial range* of the group, defined as the number of community settings inhabited during the survey year by one or more group members, was first determined. This measure was then used to calculate as a *territorial index* the percentage of all community settings (585) entered by individuals of each group. The territorial indexes of all of the different age groups are presented in Figure 2.

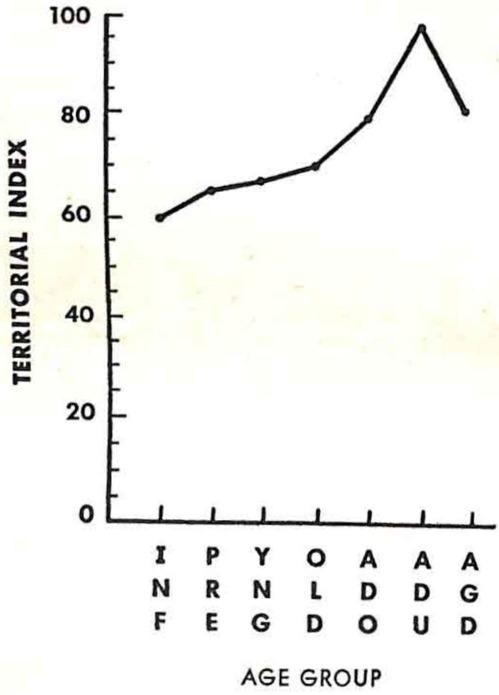


Figure 2—Percentage of all community behavior settings inhabited by members of each group.

A Midwest four-year-old once informed a station observer who had expressed astonishment at seeing him on his own far from home: "I go over the whole world." Figure 2 indicates that something like this was true for Midwest children of every age. Even for infants the territorial index was 60; and this indicator of free space rose regularly and without any sign of sudden or radical shifts of scene to 79 for adolescents.

Because territory to roam in did grow with its recipients in Midwest, children of each age level had ahead of them behavior settings beyond the occupied free area. Getting to be six when one was three, nine when one was six, and so on up promised a larger field of action. We do not know whether or how much cities differ here from Midwest. Data for comparison are available, though, from a smaller community, the Lawton home and school for crippled children (8). There, to begin with, the total number of settings was but 59. The oldest residents, 14-year-olds, entered nearly all of these settings, as shown by a territorial index for them of 95. But so it was precisely for the youngest residents, who were six-year-olds! Growing up in Lawton had little to offer. Growing up in Midwest both offered and brought continuous expansion in behavior settings.

Penetration into Settings and Age

The finding that Midwest children occupied an increasingly larger number of community behavior settings as they grew older tells nothing of how far they got into these settings and so leaves open the question as to change with age in level of involvement and responsibility in Midwest's public behavior areas.

We have rated the maximum *depth of penetration* by members of different age groups into community settings of the town on a scale that extends through six penetration zones. These zones of penetration, listed in ascending order according to their centrality or depth, and identified in terms of characteristic participation roles, with each exemplified by a single participant, are as follows: (1) *onlooker* (loafer in the Post Office); (2) *audience participant or invited guest* (spectator at a ball game); (3) *member or customer* (scout at a scout meeting); (4) *performer or active functionary* (actor in a play); (5) *joint leader* (co-proprietor of a store); (6) *single leader* (teacher of a class). The maximum depth of penetration of each age group was defined as the deepest zone reached by any member of that group during the survey year.

Figure 3 represents the distribution of the maximal depths of penetration into community settings at each of the seven age levels. It reveals an increase from infancy to adulthood and a retrogression in old age to about the adolescent level in depth of settings penetration. Again, as in expansion of territorial range, the gains made through childhood and adolescence were regular and gradual. Participations in the onlooker and guest zones (zones 1 and 2) went down while those in the member and performer zones (zones 3 and 4) went up at a slow and remarkably constant rate.

But Figure 3 points to another more conspicuous fact. It is that to stand much chance of being a leader or co-leader of any kind in Midwest you had to be an adult. Children of every age and adolescents put together penetrated to the leader zones (zones 5 and 6) in only 8 per cent of all community

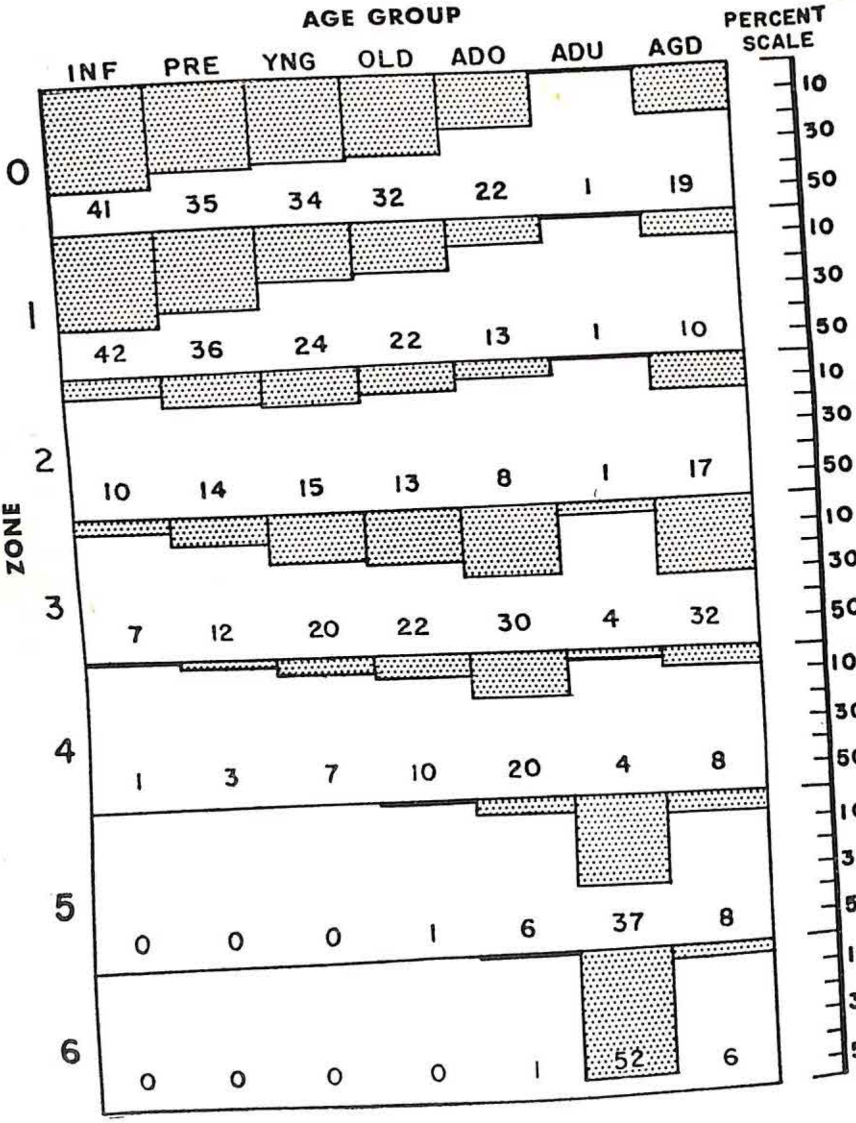


FIGURE 3—Penetration of community settings by Midwest age groups. Percentage of all community settings which age group members entered to different maximal depths of penetration. Zone 0 represents the community settings not entered by any member of an age group.

settings. Adults, by contrast, occupied these power centers in 89 per cent of all community settings.

So the freedom to roam of Midwest children was not balanced by freedom to direct. The bars in zones 5 and 6 show that, as a matter of fact, adults filled most of the leadership openings Midwest had to offer. This meant that grownups were nearly always on hand to look after children. But it meant also that children were not often left to look after themselves, with power to manage their own affairs in their own dens, clubs, gangs, or other child maintained behavior settings. Current theories are too equivocal to permit anyone to say categorically whether this was good or bad for psychological development in Midwest. It can be said that Midwest children got into less trouble than some children of cities who are reported to be "on the loose" and without adequate "supervision." Yet it is possible that they could have been better off with less adult surveillance and more independence to form and lead child groups in settings free of adults.

Figure 3 shows in any case that children in the public parts of Midwest were by no means only bystanders without status or responsibility. Older school children were at least members or customers (zone 3) in 22 per cent of all community settings. So were younger school children in 20 per cent, and even preschool children in 12 per cent. Also, 55 per cent of the settings children and adolescents penetrated to the performer zone (zone 4) were *County Teachers Institute, County Welfare Office, Women's Club I,* and other such fully "adult" centers. Children were indeed welcomed and even pressed into these settings as entertainers, clerks, office helpers, assistant mechanics, janitors, and even type setters. Our observations suggest, moreover, that for this reason the self esteem of Midwest children was high even though their power was low in most community settings.

Segregation in Settings and Age

There occurred relatively little segregation of Midwest inhabitants into narrow age, sex, social class, or race groups in community settings. But what there was of this changed some with age.

A quite regular decrease occurred from 90 in infancy to 53 in adulthood, with a reversal back to 67 in old age, in the per cent of community settings not limited in any degree (e.g., Pearl Cafe) as to age group.

Throughout childhood and adolescence, without a jog in the curve, 92 per cent of the settings entered by Midwesterners of this range were inhabited by both sexes, with this figure going down to 90 in adolescence, on down to 81 in adulthood, and back up to 84 in old age.

Ninety-seven to 98 per cent of the settings entered by children of each age group and by adolescents were inhabited by all three of the town's social groups (which correspond closely to the Upper Middle, Lower Middle, and Upper Lower categories of Warner (10)). The corresponding per cents for adults and the aged respectively were 93 and 94.

During childhood and adolescence 60 to 66 per cent of all community settings were inhabited by Negro residents (of whom there were 27) and white residents, while 54 per cent of these settings were nonsegregated as to race in both adulthood and old age.

It is clear that barriers between age groups, the sexes, the three social groups, and the two races were in no case very high in Midwest, with the exception perhaps of the race limitation. But it also is clear that age related forces tended to raise all such barriers.

Standing Behavior Patterns and Age

Differences were found from age to age in characteristics of the standing behavior patterns of community settings within the territorial ranges of the seven age groups. Demonstration of these differences rested upon ratings with suitable scales of the behavior pattern anchored within every setting.

The salient molar components of *action patterns* in community settings *entered* by infants were social contact (sociability, fellowship), recreation (pleasurable activity, games, entertainment), nutrition (dining in public places), religion (worshipping, religious study and teaching), and aesthetics (beautifying surroundings or self). A gradual shift in the relative prominence of different action patterns occurred through childhood to adulthood to old age, where the most prominent patterns were social contact, earning a living (hiring out for wages), recreation and government (creating, interpreting, and enforcing laws and regulations).

The molecular components or behavior mechanisms outstanding in community settings entered by infants were talking (verbalization of all kinds, including singing and yelling), affective behavior (overtly expressed emotionality), and thinking (decision making and problem solving). Again, a process of gradual change took place until, in old age, the most prominent mechanisms in the inhabited settings were thinking, gross motor activity, and manipulation (use of the hands in prehension).

Both of these courses of change, the one in goal directed action, and the other in act mechanisms, point to a movement from relatively little involvement, during the earliest years, to relatively great involvement, in old age, with necessities of the real world.

DEVELOPMENTAL CHANGE IN BEHAVIOR EPISODES

Episodes as Individual Behavior Units

Psychological development in Midwest, while it proceeds in behavior settings as community units, necessarily involves all of the time units of behavior and situation in the lives of individual persons. Here also the Midwest behavioral panorama includes readily observed targets of description, which we have called *behavior episodes*. These have been used as standard ecological units to take into account the behavior and psychological habitat of individual Midwest children.

To collect episodes for study we have employed the *specimen record,* a detailed, sequential, narrative account by skilled observers of an individual child's behavior through a more or less extended time. At its longest a record of this kind spans a day—and some 300 pages. Here is a brief sequence from a specimen day record on Mary Ennis, an eight-year-old girl. It occurs in the Ennis home, to which Mary has just returned from school with a box of paints and other belongings. Mary stands in her bedroom, from which one door leads to a dining room, and another to a bathroom. Her mother is seated in the dining room, and the observer stands across the bedroom from Mary.

Painting Lips

Mary took her box of paints and went directly and purposefully into the bathroom.

She turned on the water.

A mirror was behind the door and because the door was open a little, I could see through the crack that Mary was standing at the mirror.

Mary said, "I'm trying to do something," and giggled a very mischievous giggle.

Her mother evidently recognized the quality of the giggle because she came into the bedroom to see what was going on.

We both saw that Mary was painting her lips with the paint brush and the water color paint.

Washing Paint Off Face

Mrs. Ennis said, "Oh, you had better not do that. That will make your lips chapped." She added for good measure, "I don't think it will come off," and then left the room.

Immediately, almost, Mary stepped from the mirror to the sink. She started to run the water and wash off the paint.

Challenging Mother

She said, "What did you say? Did you say it wouldn't come off?" This was a challenge to the mother, which Mary repeated with a good deal of emphasis, saying, "Mommie, you said it wouldn't come off and it did."

The mother answered, "I didn't think it would." She wasn't arguing, just explaining.

Mary worked still at the sink. She turned on the water full force and let it run for a full minute while she washed off the brush and fussed around with the paints.

Then she turned the water off and came back into her bedroom.

Readers will have noted that this sequence is marked off by side brackets into three parts, entitled: *Painting Lips, Washing Paint Off Face,* and *Challenging Mother.* These we take to be episodes. Other episodes of our records include *Working Problem at Board, Riding Father's Foot, Making Rock Garden, Getting Splinter Out of Hand, "Painting" Porch with Mud,* to which I will return later, *Telling Off Friend, Sweeping Floor,* and *Getting Dressed.*

Every episode contains an action and a particular situation that correspond respectively to the standing behavior pattern and the milieu context of a behavior setting. The action is always a segment of behavior with a constant direction, and the situation is the coexisting set of conditions in the person and the environment which interact to determine the action. Take *Washing Paint Off Face,* of Mary's day. This title identifies the action and indicates its direction. The situation includes on the person side the need state implied by Mary's going on her own to the sink. It indicates on the environment side the cautioning mother, the warning she expressed, the familiar rooms, the mirror, the paints, the sink, and the water as carriers of psychological effects. There are as parts of the situation, in short, all of the conditions that made a difference to the behavior of this child at the time, as these are represented in the record. From this much it will be understood further that an episode is a *molar* behavior phenomenon as defined by Tolman (9) and Lewin (6, 7).

Conceptual and operational criteria of episodes have been developed in some detail (3, pp. 225-270). These criteria have been tested with resulting evidence that the process of discriminating episodes in specimen records is reliable (3, pp. 270-273).

Recorded episodes can be studied with sole regard for their action and situation content. But they also can be studied from a quite different point of view. To divide a specimen record into episodes is to find parts of a whole behavior sequence. Inquiry can then be made into relations of part to whole and part to part within the sequence without any reference to what the parts contain. This, in the terminology we have used, is to ask about the structure or anatomy of behavior. We have analyzed the episodes in records on Midwest children from the standpoint of both their action and situation content and their structural properties, with consequent findings in each case on change in behavior and habitat with age. I will try, therefore, to represent here each of the two types of analysis, beginning with the anatomy problem.

Behavior Structure

The operation of episoding opens this problem to study by indicating on the margin of a specimen record some basic whole to part and part to part relations. These can be reviewed quickly.

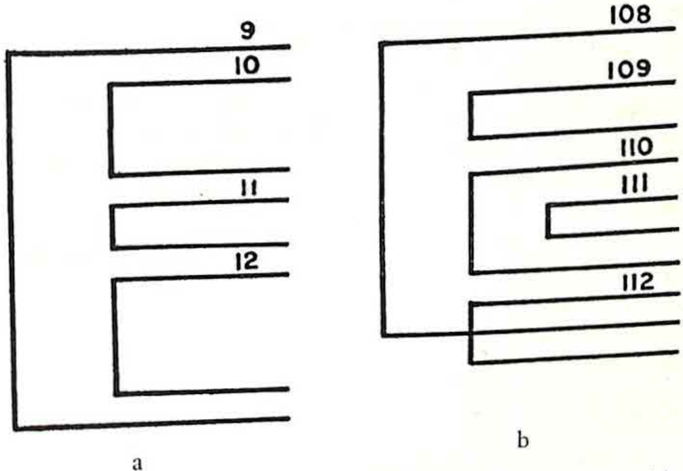


FIGURE 4—a: An episode linkage from a day of a Midwest two-year-old.
b: An episode linkage from a day of a Midwest ten-year old.

First, different episodes can occur at one time, in which case they are said to *overlap*. Overlapping is graphed by parallel placement of episode brackets. The brackets that bound *Washing Paint Off Face* and *Challenging Mother* of Mary's day, for example, are so placed to show overlap such that the whole of the shorter unit occurred while a part of the longer one was in progress. As many as four episodes, though rarely more than three, have been found in progress at one time in our records. Three were judged to overlap, for example, while Wilma sat eating toast, petting her dog, and asking if she could go outdoors to play. Where, as in the case of Mary's *Painting Lips,* a single episode occurs alone, we say it is *isolated*.

Any combination of episodes which overlap to form an interwoven continuum we have called a *linkage*. The sequence from Mary's day extends through two linkages of which the first has only one *link*. Different episodes of a linkage often overlap sequentially with a single continuously ongoing unit. See Figure 4a. It represents only by episode brackets a part of a day record on two-year-old Mary Chaco. At 7:31 A.M., the ninth unit of the day began. Then, while it continued, Mary became involved successively in episodes 10, 11, and 12. A linkage can be much longer and more complex. Figure 4b graphs a just moderately intricate one, presenting unit to unit relations which readers themselves may note, from a day record on 10-year-old Douglas Crawford.

A sketch of our main findings on behavior structure will provide a basis for description of differences age made in this area.

The number of episodes per day varied from approximately 400 to about 1,300 through a median of 850 (after correction for slight inequalities in length of day). These episodes ranged in duration from a few seconds to 121 minutes through a median of 45 seconds, with 70 per cent shorter than two minutes. Linkages averaged about 400 per day. These sequences of structurally interrelated units included as many as 81 episodes, with the median number 2.3; and their average duration was about two minutes. Fifty to 75 per cent of the episodes per record overlapped with at least one other. Only 10 per cent of these concurrent units, however, overlapped with more than two others. Roughly 11 per cent of the episodes per record were left incomplete, after nonattainment, failure, frustration, or interruption from without. Two per cent or less in each day were broken off and then resumed later. When this happened, the number of interpolated episodes varied greatly, but made up no more than 10 per cent of the units in any record.

TABLE 3

RANK ORDER CORRELATIONS BETWEEN AGE AND VARIABLES OF BEHAVIOR STRUCTURE

Area	*Variable*	*tau*	*p*
Unit differentiation			
	Number of episodes	−.61	.01
	Number of linkages	−.67	.01
	Episodes per linkage	.73	.01
Unit duration			
	Minutes per episode	.61	.01
	Minutes per linkage	.67	.01
Position factors			
	Episodes overlapping	.61	.01
	Episodes overlapping with five or more others	.50	.03
	Episodes between interrruptions and resumptions	.42	.07
Episode termination			
	Episodes left incomplete	−.74	.01
	Episodes ending with another ongoing	.64	.01

NOTE.—In this table and hereinafter the coefficient is *tau* and the *p* test is two-tailed.

The rank order correlation coefficients of Table 3 sum up our findings on change in behavior structure with age. These coefficients show or at least suggest that psychological correlates of increasing age brought about for the older children as compared with the younger: fewer but longer episodes; fewer but longer and also more complex linkages of episodes; more frequent occurrences of different episodes at the same time; fewer incomplete episodes; and more episodes between the interruptions and the resumptions of discontinuous episodes—with the last implying, perhaps, greater ability to sustain need tensions.

These data are consistent with the expectation of greater complexity in the behavior of older children and with commonly mentioned marks of maturity. So it appears especially from the findings that the younger children tended more than the older to do things sequentially, one at a time, to move often from one action to another, and to leave undone what they started, while the older children tended more than the younger to get involved in more than one thing at a time, to engage in lasting actions, and to complete what they set out to do.

The results here open speculation about their meaning for the psychological ecology of Midwest as compared with other communities. Comparable results from children on isolated farms, say, or from children of the Loop in Chicago could have been quite different. Complex overlapping of episodes or extended interlinking of episodes, for example, may occur less often on the farm, where behavior settings are scarcer, reputably more simple, and less closely jammed together. Both of these complications, on the other hand, may occur more often in the Chicago Loop. Such differences might conceivably be used, moreover, as partial indexes of more profound dissimilarities in the tempo, the complexity, and perhaps the stresses and strains of community life. This can be only guessing now. But it has seemed to us that anything as intelligibly sensitive to age differences as behavior structure appears to be must also be sensitive in some ways to culture and community differences.

There remains a larger outcome of this exploring which may have potentialities for research on general principles of psychological development. It is simply demonstration that the psychological continuum has fundamental properties which can be described reliably in terms of behavior structure. How many things does a child do in a day? How are these things related as continuum units? We do not know that such questions have been asked earlier. But our present point is only that these questions evidently can be answered in a reliable way. To be added here to the earlier mentioned evidence that specimen records can be episoded reliably there is evidence that application of the present descriptive categories also is reliable (2. pp. 274-285). That the data have internal consistency is pertinent as well in this connection. So finally are indications of agreement between the data and theories resting on independent and quite different information. This holds especially for what we have taken to be affinities (discussed in 2, pp. 300-302) between our structure findings and developmental theories of Lewin (7), Kounin (5), and Baldwin (1).

Psychosocial Habitat and Behavior

Episodes are far from being the neutral elements which their structural analysis as parts of the behavior continuum might suggest. Two ways in which this is true are important here. First, psychological development occurs in and through episodes. Second, episodes often are arenas for face to face social interaction.

We have used episodes in specimen records on both of these counts in a study of the psychosocial habitat and behavior of Midwest children. Our data indicate that 60 to 80 per cent of these action units involve the children in relationships with mothers, fathers, teachers, peer friends, neighbors, pets, and others of Midwest. What are the properties of these associates as carriers of psychological effects? This is chiefly a question of how others behave in relation to the children. But how, also, do the children behave in relation to others? And what kinds or patterns of interaction per se occur as between the children and their associates?

We have applied to recorded episodes of Midwest children a set of categories dealing with elemental aspects of social action and situation to get some answers to these questions. An example may be helpful. The mother of Margaret Reid, a Midwest four-year-old, orders Margaret to stop "painting" a porch pillar of her house with mud. Margaret says, "I want to paint." This interaction continues and gets increasingly complicated in the course of a troubled episode lasting six minutes. The following are the main questions asked about this episode, as about others, by means of the categories at hand.

As between the mother and Margaret, what is the ratio of **power**? In other words, as Margaret saw it, who of the two has the greater potential ability to change the behavior of the other? A rating is made on an 11-point scale, with equal power at its center.

What degree of "used power" or **social pressure** is the mother actually bringing to bear on Margaret, and vice versa. A rating is made here on a 5-point scale.

What **mode of action** does each of these interacting persons exhibit? *Dominance,* defined as authoritarian pressure? *Aggression,* defined as direct physical or verbal attack? *Submission,* to either dominance or aggression? *Resistance,* against either of these? *Appeal,* for help, a gift, or protection? *Nurturance,* in the sense of acting to benefit the other person by extending help, a gift, or protection? *Avoidance,* in the sense of hanging or drawing back or withdrawing from, the other person? Presence or absence is judged on every action mode of these seven. Also, the same judgment is made on *compliance,* defined as doing what the other person wants as a response to appeal rather than to authoritarian pressure.

What degree of momentary positive or negative *affection* does Margaret show for her mother, and the mother for Margaret? The same is asked about *evaluation* in the form of approval or disapproval by each of the other's conduct. And what is the *mood* of each in terms of degree of happiness or unhappiness? Each of these three **attributes of action**—affection, evaluation, and mood—is rated on an 11-point scale, with neutral at its center.

Some secondary questions also are raised.

In what **behavior setting** does the episode occur? The answer here: Home Outdoors.

Is the **action circuit** *open* or *closed*? Is there action, in other words, by only one of the associated persons, or by both? Margaret responded to the mother, so the answer here is *closed.*

What level of **associate complexity** is there on the associate side, *simple* or *compound*? If simple, the child is acting in relation to one person; if compound, in relation to two of more persons, as when a pupil reports to a class. In the present case, the mother made up the whole of Margaret's social behavior object, which means that the answer here is *simple*.

Who took the **initiative** in the transaction? The associate or the child? Here, obviously, the associate.

These and ancillary categories, which include some to identify both subject and associate, plus others to describe the target unit in terms of its structural properties, we have applied wherever each was applicable to approximately 10,000 episodes of specimen day records on four boys and four girls differing in age from two years through ten. Tests made on representative samples of these episodes have shown varying but generally satisfactory agreement between independent analysts in application of the categories. As in reporting the results on behavior structure, I will start here with a partial summary of overall findings, and then present some data on age relationships.

Sixty to 80 per cent of the episodes in every day involved the children in relations with an associate. A complete social interaction, such that the child and the other person each acted in relation to the other, occurred in approximately four-fifths of these episodes. Associates of the children usually were single individuals rather than groupings of two or more individuals.

Adults far exceeded other children as associates. The most nearly omnipresent among these adults were mothers, who figured in about three times as many episodes as fathers. Teachers were second only to mothers in the number of episodes they entered. As these findings suggest, greater rated power of associates was quite generally characteristic of the children's social situations. Adults, especially, greatly exceeded the children in power. The children nonetheless initiated approximately half of their interactions with individuals of every associate grouping.

The data on modes of action are of special interest here. Figure 5 represents the frequency of each mode in action by, and in action of the children toward, both adult associates and child associates. All of the different modes are identified by prototypic verbal expressions.

Two complementary relationships appear in the mode findings for episodes with adult associates. First, a *control* relation is indicated by the evidence that adults were often dominating while, at the same time, the children were often submissive. Second, a *dependency* relation is indicated by the evidence that children often appealed to adults while adults were often nurturant. These two relationships, the one between "bossing" and giving in, and the other between asking for benefit and getting it, can be brought together in the conclusion that the social regimen of the children with adults was a kind of *benevolent autocracy*. This regimen had, further, the common authoritarian feature of power centralization. As data which cannot be added here show, the power was centered preeminently in mothers and teachers; the autocracy was matriarchal. Readers may find, however,

	ADULT ASSOCIATES		CHILD ASSOCIATES	
Mode of Action	*Associate Side*	*Child Side*	*Associate Side*	*Child Side*
Dominance:				
Quiet! Quiet!	IIIIIIIIIIIIIIIII	IIII	IIIIIII	IIIIIIIIIIII
Nurturance:				
Let me help you	IIIIIIIIIIIII	IIII	IIIIIIIIII	IIIIIIIII
Resistance:				
No	IIIIII	IIIIIIII	IIIIIIIIIIII	IIIIIIII
Appeal:				
Please	IIIII	IIIIIIIIIIIIIII	IIIIII	IIIIIIIII
Compliance:				
I will	IIIII	IIIIII	III	III
Aggression:				
Take that!	III	III	IIII	IIIIII
Submission:				
I give in	II	IIIIIIIIIII	IIIIII	III
Avoidance:				
You repel me		I	II	I

FIGURE 5—Modes of action by Midwest children and their adult and child associates. Each row of symbols represents the number of ocurrences of a mode in every 50 episodes. For example, the row of symbols on the child side, under adult associates and opposite dominance, shows that in four episodes of every 50 dominance occurred in behavior of the children with adult associates.

as we have concluded by studying Figure 5 and related data, that this adult with child situation was not characteristically subjugative or bridling (note high child resistance) or hostile (note high adult nurturance, low adult aggression) or rejective (note zero adult avoidance) or conflictive (note low counter dominance on the child side, low aggression on both sides).

Figure 5 shows that the social pattern of children with children was not featured by an authoritarian control relation. Thus, although child associates were often dominative, the children were not at all equally submissive. The subjects tended more to be resisting, and they were even more dominative than their child associates. The possibility of a transfer to weaker individuals of impulses to oppose adult authority is suggested by relatively more frequent dominance, aggression, and resistance by the children against children than by the children against adults.

Data from the *mood, affection,* and *evaluation* categories indicate that the social situations of the children were preponderantly pleasant, friendly, and approving, in line with the evidence of frequent nurturance from both adults and other children. Also, findings on *kind of interplay* indicate that interactions of the children with others were harmonious much more often than disharmonious, although outright conflict, discord short of conflict, and unfriendly rivalry did occur in these interactions with considerable frequency. One apparent reason for this is that, as the data on relative power and the indications of authoritarian control by adults lead one to anticipate, *social pressure* by others on the children greatly surpassed in frequency and strength social pressure by the children on others.

Table 4 presents coefficients which indicate a number of changes with increase in age of the children in these various aspects of social behavior and situation. The more prominent of these changes follow.

TABLE 4

RANK ORDER CORRELATIONS BETWEEN AGE AND PSYCHOSOCIAL VARIABLES

Area	*Variable*	*tau*	*p*
Associate identity			
	Episodes with adults	—.79	.01
	Episodes with mothers	—.71	.05
	Episodes with fathers	—.64	.05
	Episodes with children	.64	.05
	Episodes with compound associates	.82	.01
Social habitat			
	Average power of all associates	—.93	.01
	Average power of child associates	—.71	.05
	Average fluctuation in relative power	.57	.10
	Frequency of appeal by all associates	.82	.01
	Frequency of appeal by adults	.68	.05
	Frequency of resistance by adults	—.61	.05
	Frequency of resistance by mothers	—.71	.05
	Frequency of avoidance by adults	—.71	.05
	Average plus evaluation by fathers	—.62	.10
Social behavior			
	Frequency of appeal to all associates	—.61	.05
	Frequency of compliance with all associates	.67	.05
	Frequency of compliance with adults	.71	.05
	Frequency of nurturance of adults	.64	.05
	Frequency of nurturance of children	.67	.05
	Average plus evaluation of fathers	—.62	.10
	Average plus affection for children	—.60	.05
Social interaction			
	Frequency of conflict with children	—1.00	..
	Frequency of conflict with adults	—.57	.10

Other children waxed while adults in general and parents in particular waned as associates of the children. Associate groupings of two or more persons increased. Relative power of the children with associates en masse rose. So did fluctuations from episode to episode in relative power, as the ratio of persons younger than the child to persons older than the child necessarily increased. Appeals to the children went up in frequency while resistance against them and avoidance of them went down. On the child side, appeals to others went down while nurturance of others and compliance with others went up. Conflict decreased. Yet positive evaluation fell on both sides in relations of the children with their fathers; and affection of the children for child associates declined.

The children evidently did not outgrow their control relation with adult associates. In particular, neither dominance by adults nor submission to them by children declined. Some loosening of adult control is suggested by the decrease in resistance by adults. But the basic control relation evidently persisted. And so also did the benevolent side of the dependency relation with adults, inasmuch as grownups nurtured the older children no less than the younger. The older children, though, appealed to adults less, and nurtured them more. It is as if grownups persisted in saying as the children gained in size and strength, "You still need us as much as ever," while the children said, "We need you less" and "You need us more."

These relationships agree generally with sense and one another. They contribute some intelligible measures of increasing social maturity and of compatible change in naturally occurring conditions of the social environment.

Discussion

Biological ecologists have prodigious knowledge of how fish grow in streams, trees in forests, and the like. But there have been virtually no psychological ecologists, with the result that we have known little about the psychological growth of persons in communities. The data here add some knowledge of this kind from the one community of Midwest.

Animal and plant ecologists all have had to adopt standard units of description, such as *grazing ranges, behavior cycles, successions,* and *life cycles*. This need we have tried to meet by choosing as descriptive units behavior settings and episodes. Our first claim for these targets of a psychological ecology would be that description in their terms can take into account both psychological events and conditions under which these same events occur for the reason that each target is a unit of behavior and its context.

Our first claim for the present data, too, would be that they describe developmental change in context or situation as well as in behavior. Procedures in laboratories, testing rooms, and clinics have produced most of our scientific knowledge about psychological growth. These procedures have taught us much about change in behavior with development, as witness scores of age curves on motor behavior, language behavior, learning behavior, emotional behavior, social behavior. But age curves on change in psychological situation with development are rare indeed. The main reason for this we take to be that the situations in which behavioral growth occurs cannot often be brought into laboratories, testing rooms, or clinics, but generally must instead be left in the field, where they can be studied only by field methods. Some analogues of naturally occurring situations can, of course, be arranged and observed, as rivers are simulated and their laws explored, by experiments that manipulate what they study. This has worked well in psychology. Yet, in the first place, some naturally occurring psychological conditions are like some conditions of weather cycles, ocean currents, and bird migrations in that they cannot be arranged, easily or at all; and, in the second place, such experimental investigation can never solve the purely descriptive, ecological problems which seem to us important in their own right, as they have been in other biological and social sciences.

Also to be considered here is the prospect that psychological field methods can be used beyond descriptive study for inquiry into interdependencies of the sort generally associated with laboratory research. With the present methods, for example, one can explore relations within settings, within episodes, or across the boundaries of settings and episodes. This has had to wait until now on the prior task of straight description with its continuing problems of method. But some such line of investigation via analytical field methods of one kind or another is encouraged in principle by the accomplishments of public health researchers, agronomists, geologists, anthropologists, astronomers, and other descriptive scientists who regularly educe basic statements of interdependence by interrelating naturalistic data.

This leaves implicit in the main idea of psychological ecology, research on behavior and its conditions in differing communities of different cultures. Experiments in nature on consequences for psychological habitat and behavior of differences between communities in size, independence, completeness, population density, rural or urban character, suburban locale, technological advancement, social organization, and various culture dimensions are logical possibilities of such comparative investigation. Whether conceptual and technical requirements of these experiments can be met is an important question which the present methods and data, added to those of earlier field studies in social science, may help to answer.

REFERENCES

1. BALDWIN, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.
2. BARKER, R. G., & WRIGHT, H. F. *One boy's day.* New York: Harper, 1951.

3. Barker, R. G., & Wright, H. F. *Midwest and its children.* Evanston: Row, Peterson, 1955.

4. Bates, M. *The nature of natural history.* New York: Scribner, 1950.

5. Kounin, J. S. Intellectual development and rigidity. In R. G. Barker, J. S. Kounin, & H. F. Wright (Eds.), *Child behavior and development.* New York: McGraw-Hill, 1943.

6. Lewin, K. *Principles of topological psychology.* New York: McGraw-Hill, 1936.

7. Lewin, K. *Field theory in social science.* New York: Harper, 1951.

8. Newton, M. R. A study in psychological ecology: the behavior settings in an institution for handicapped children. Unpublished master's thesis, Univer. of Kansas, 1953.

9. Tolman, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century, 1932.

10. Warner, W. L., Meeker, Marchia, & Eells, K. *Social class in America.* Chicago: Science Research Associates, 1949.

11. Wright, H. F., & Barker, R. G. *Methods in psychological ecology.* Lawrence, Kansas: Department of Psychology, Univer. of Kansas, 1950.

12. Wright, H. F., Barker, R. G., Nall, J., & Schoggen, P. H. Toward a psychological ecology of the classroom. *J. educ. Res.,* 1951, 45, 187-200.

CHILD DEVELOPMENT

Vol. 27
September, 1956
No. 3

Editor
WILLIAM E. MARTIN
Purdue University

Consulting Editors

REGINALD MACGREGOR ARCHIBALD
Rockefeller Institute for Medical Research

JOSEPH H. BRITTON
Pennsylvania State University

SIDNEY W. BIJOU
University of Washington

VAUGHN J. CRANDALL
Fels Research Institute

ERCEL S. EPPRIGHT
Iowa State College

C. GLEN KING
Nutrition Foundation

BENJAMIN PASAMANICK
Ohio State University

ANSELM STRAUSS
University of Chicago

RALPH L. WITHERSPOON
Florida State University

Managing Editor
KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Bureau Ed[illegible] Research
DAVID H[illegible]
[illegible]ING CO[illegible]
Dated
Accs. No

NOTICE TO SUBSCRIBERS

For all new subscriptions and renewals for 1957 and succeeding years received after September 15, 1956, the following rates will apply: CHILD DEVELOPMENT—$8.00 domestic; $8.50 foreign; single issues, $2.50. CHILD DEVELOPMENT ABSTRACTS & BIBLIOGRAPHY—$6.00 domestic; $6.50 foreign; single issues, $2.50. MONOGRAPHS of the SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT—$6.00 domestic; $6.50 foreign; single issue, price varible. CLUB SUBSCRIPTION to all three publications—$16.00 domestic; $17.50 foreign.

Printed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

CHILDREN'S WORK ATTITUDES AND SIBLING CHARACTERISTICS [1,2]

HELEN L. KOCH
University of Chicago

PROBLEM AND PROCEDURES

This study is part of a larger one concerned with the effects on the personality of children of characteristics of their siblings. We are reporting here our data bearing on the relation between certain work habits and attitudes of the five- and six-year-old from two-child, white, native-born, intact, urban families and the ordinal position of the child, the age difference which separates him from his sibling and his sex in relation to that of his sibling. Sex differences were also noted. The habits and attitudes explored were aggressiveness or initiative, curiosity, originality, number of interests, enthusiasm, ambition, responsibleness, tenacity, planfulness, and tendency to dawdle or procrastinate.

Our child subjects numbered 384. The experimental design included three sib-spacing levels, two ordinal positions, subjects of two sexes and siblings of two sexes. There were 48 children in each of the following categories—male with a male sib older, male with a male sib younger, male with a female sib older, male with a female sib younger, female with a male sib older, female with a male sib younger, female with a female sib older and female with a female sib younger. Each of these groups of 48 children was composed of three subgroups of 16 children, representing the following three sibling-age-difference levels: siblings differed in age by under two years, by two to four years, and four to six years, respectively. Hence our basic subgroups of 16 each numbered 24.

---

[1] The investigation was supported, in part, by a grant from the National Institute of Mental Health, of the National Institutes of Health, U. S. Public Health Service, and, in part, by the Sonia Shankman Orthogenic School Research Fund, University of Chicago.

[2] The author wishes to acknowledge the help received from, and to express her gratitude to, the teachers and principals of the following schools in Chicago which participated in the Study: Bennett School, Bradwell School, Field School, Holmes School, Gunsalus School, Luella School, Perry School, Schmid School, Sullivan School, Wentworth School, University of Chicago Laboratory School and the University of Chicago Nursery School. The author is also much indebted to Dr. Grace Munson and Dr. Thelma Thurstone, both formerly Directors of the Bureau of Child Study, Chicago Public Schools. Dr. Helen Heath did most of the statistical work for the study. We acknowledge with gratitude her discriminating and loyal services.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

TABLE I

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: *AGGRESSIVENESS** AND *CURIOSITY**

SUBGROUP Subject Sex	Ordinal Position	Sibling Sex	*Aggressiveness** 7–24	25–48	49–72	Total	*Curiosity** 7–24	25–48	49–72	Total
			AGE DIFFERENCE BETWEEN SIBLINGS IN MONTHS							
			SCORE MEANS				SCORE MEANS			
M	2	M	3.926	4.068	4.471	4.155	3.699	3.728	4.284	3.904
M	2	F	3.782	3.793	3.860	3.812	3.854	3.848	3.942	3.881
F	2	M	3.982	4.157	4.114	4.084	3.774	4.094	3.827	3.899
F	2	F	3.990	4.131	4.047	4.056	3.966	3.604	3.913	3.828
M	1	M	3.646	3.993	4.313	3.984	4.179	3.847	4.211	4.079
M	1	F	3.580	4.616	3.933	4.043	3.906	4.344	4.448	4.233
F	1	M	4.392	4.114	4.466	4.324	4.378	4.213	4.552	4.381
F	1	F	4.154	4.061	3.924	4.046	3.876	3.815	4.133	3.941
M	2	M or F	3.854	3.930	4.166	3.983	3.776	3.788	4.113	3.892
F	2	M or F	3.986	4.144	4.081	4.070	3.870	3.849	3.870	3.863
M	1	M or F	3.613	4.304	4.123	4.013	4.043	4.095	4.329	4.156
F	1	M or F	4.273	4.088	4.195	4.185	4.127	4.014	4.342	4.161
M or F	2	M	3.954	4.112	4.293	4.120	3.737	3.911	4.056	3.901
M or F	2	F	3.886	3.962	3.953	3.934	3.910	3.726	3.928	3.854
M or F	1	M	4.019	4.053	4.390	4.154	4.278	4.030	4.381	4.230
M or F	1	F	3.867	4.338	3.928	4.045	3.891	4.079	4.290	4.087
M	1 or 2	M	3.786	4.030	4.392	4.070	3.939	3.788	4.248	3.991
M	1 or 2	F	3.681	4.205	3.896	3.927	3.880	4.096	4.195	4.057
F	1 or 2	M	4.187	4.136	4.290	4.204	4.076	4.153	4.189	4.140
F	1 or 2	F	4.072	4.096	3.985	4.051	3.921	3.709	4.023	3.884
M	1 or 2	M or F	3.734	4.117	4.144	3.998	3.909	3.942	4.221	4.024
F	1 or 2	M or F	4.129	4.116	4.138	4.128	3.998	3.931	4.106	4.012
M or F	1 or 2	M	3.987	4.083	4.341	4.137	4.008	3.971	4.218	4.066
M or F	1 or 2	F	3.876	4.150	3.941	3.989	3.900	3.903	4.109	3.971
M or F	2	M or F	3.920	4.037	4.123	4.027	3.823	3.819	3.992	3.878
M or F	1	M or F	3.943	4.196	4.159	4.099	4.085	4.055	4.336	4.158
M or F	1 or 2	M or F	3.932	4.117	4.141	4.063	3.954	3.937	4.164	4.018

We shall refer to these groups in code as follows: the first letter and number will indicate the sex of the child subject and his ordinal position, the second letter the sex of the sibling and the numbers in parentheses the age difference in months which separates the child and his sib. Thus M2F(25-48) refers to second-born boys who have one sister two to four years older.

The measures of the dependent variables were teachers' ratings. Five of the scales (the starred ones in Tables 1 to 5) were taken from the Fels Child Behavior Rating Scales (16); five (the unstarred scales in Tables 1 to 5), from the California Behavior Inventory for Nursery School Children (4). The ratings were made by the teachers on a line scale. Their checkings were then transmuted to ratings on a nine-point scale. These ratings, in turn, were normalized on the basis of a population of 498 five- and six-year olds of which our subjects were the major part. Four was taken as the mean score for each trait in the total population. The Bartlett test (7, pp. 195-200) was applied to determine whether the ratings could be assumed to be drawn from populations with similar variance and this was found to be the case. For the likely reliabilities of our measures, see (5) and (16).

The means for our various groups are presented in Tables 1 to 5 and the analysis of variance data are offered in Table 6. Only data for the total population are presented. We shall report in the text significant relations that obtain only for subgroups. In Table 6, *F*'s are included only when the group differences are significant at the 5 per cent point or better. When it seemed indicated, *t*'s were computed for the subgroup differences but in the interest of economy these are not presented. It can be assumed the differences we discuss are significant, or very near significance, at the 5 per cent point or better unless we specify otherwise.

The groups were matched individual by individual in age and in the socio-economic status of the father's occupation (8) as well as of the neighborhood of the family residence (14). The group compositions have been described in detail in an earlier publication in this journal (11). Matching, of course, could not be exact, but 98 per cent of the children were matched within six months of age, and 75 per cent within four months; 93 per cent, within one occupational level; and 88 per cent, within one rank on the neighborhood scale. Level 4 of the occupation scale was omitted, as our subjects were all city children. The children were drawn chiefly from the public schools of Chicago, only one private school system, the University of Chicago Laboratory Schools, contributing subjects. (For a discussion of some of the matching and sampling problems we faced the reader is referred to (10) and (11).)

The traits studied probably, in the main, need no definition here. The trait, aggressiveness, refers to initiative and dominance rather than to hostility. By tenacity is meant persistence toward a goal and long attention span.

TABLE 2

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: *ORIGINALITY** AND *NUMBER OF INTERESTS*

SUBGROUP			*Originality**				*Number of Interests*			
Subject		Sibling	AGE DIFFERENCE BETWEEN SIBLINGS IN MONTHS							
Sex	*Ordinal Position*	*Sex*	*7–24*	*25–48*	*49–72*	*Total*	*7–24*	*25–48*	*49–72*	*Total*
			SCORE MEANS				SCORE MEANS			
M	2	M	3.748	4.085	4.204	4.013	3.740	4.079	4.491	4.103
M	2	F	3.814	3.906	3.786	3.835	3.630	3.807	3.782	3.740
F	2	M	3.976	4.326	4.156	4.152	3.840	3.979	4.091	3.970
F	2	F	4.309	3.778	4.111	4.066	4.175	3.882	3.886	3.981
M	1	M	3.860	4.009	4.006	3.958	3.878	3.846	4.074	3.933
M	1	F	4.063	4.072	4.043	4.028	3.814	4.080	4.405	4.100
F	1	M	4.144	4.096	4.164	4.135	4.132	4.134	4.344	4.203
F	1	F	3.961	3.937	3.817	3.905	3.944	3.734	4.166	3.948
M	2	M or F	3.781	3.995	3.995	3.924	3.685	3.943	4.137	3.922
F	2	M or F	4.143	4.052	4.133	4.109	4.008	3.931	3.988	3.976
M	1	M or F	3.962	3.993	4.024	3.993	3.846	3.963	4.239	4.016
F	1	M or F	4.052	4.017	3.991	4.020	4.038	3.934	4.255	4.076
M or F	2	M	3.862	4.205	4.180	4.082	3.790	4.029	4.291	4.037
M or F	2	F	4.062	3.842	3.948	3.951	3.903	3.844	3.834	3.860
M or F	1	M	4.002	4.053	4.085	4.047	4.005	3.990	4.209	4.068
M or F	1	F	4.012	3.957	3.930	3.966	3.879	3.907	4.285	4.024
M	1 or 2	M	3.804	4.047	4.105	3.985	3.809	3.963	4.283	4.018
M	1 or 2	F	3.938	3.941	3.914	3.931	3.722	3.944	4.093	3.920
F	1 or 2	M	4.060	4.211	4.160	4.144	3.986	4.057	4.217	4.087
F	1 or 2	F	4.135	3.858	3.964	3.985	4.059	3.808	4.026	3.964
M	1 or 2	M or F	3.871	3.994	4.010	3.958	3.766	3.953	4.188	3.969
F	1 or 2	M or F	4.097	4.034	4.062	4.065	4.023	3.932	4.122	4.026
M or F	1 or 2	M	3.932	4.129	4.133	4.064	3.898	4.010	4.250	4.052
M or F	1 or 2	F	4.037	3.899	3.939	3.958	3.891	3.876	4.060	3.942
M or F	2	M or F	3.962	4.024	4.064	4.017	3.846	3.937	4.063	3.949
M or F	1	M or F	4.007	4.005	4.008	4.006	3.942	3.949	4.247	4.046
M or F	1 or 2	M or F	3.984	4.014	4.036	4.011	3.894	3.943	4.155	3.997

RESULTS

We shall here summarize the major findings of the study in outline form, since they are rather complicated, many interactions being significant. Discussion will be reserved for a later section.

Sib-Age-Disparity-Group Differences

1. The wider the age gap between the sibs the higher were the ratings on number of interests (Tables 2 and 6).

2. At the middle spacing FM's exceeded FF's in rating on tenacity and at the close spacing F1M's scored higher than F1F's but M2M(49-72)'s exceeded in score M2F(49-72)'s (Tables 4 and 6).

3. At the close spacing girls received a higher mean rating in aggressiveness than boys (Tables 1 and 6).

4. At the widest spacing those with a male sib were judged more aggressive than those with a female sib (Tables 1 and 6).

5. At the sib-age-difference levels under four years girls were gauged more tenacious than boys (Tables 4 and 6). Sex differences were greatest among the children who had a brother.

Sib's-Sex-Group Differences (See Items 2 and 4.)

6. The children with a brother were assessed as more aggressive, on the average, than those with a sister, when the sib disparity in age was four to six years (Tables 1 and 6).

7. Those children from opposite-sex sib pairs who differed from their sib two to four years in age were rated higher in curiosity than those from same-sex pairs. F1M's at all spacings scored higher than F1F's (Tables 1 and 6).

8. The children among first-borns whose sib differed in sex received higher ratings on enthusiasm and number of interests than the children whose sib was of the same sex; but among second-borns those with a male sib exceeded in their rating on enthusiasm those with a female sib and M2M's scored higher than M2F's in number of interests (Tables 2, 3 and 6).

9. In the main, the ratings on ambition of the children with a male sib were higher than those for children with a female sib. This trend was especially marked in the case of girls at the spacings under four years and in the case of second-born males (Tables 3 and 6).

10. First-born girls with a brother at the sib spacings up to four years received higher ratings on tenaciousness or dominance than did those with a sister, and at the middle spacing the sib-sex relation held among second-born girls as well. When the sibs differed in age four to six years, M2M's scored higher than M2F's (Tables 4 and 6).

Ordinal-Position-Group Differences (See Items 2, 8 and 10.)

11. First-borns scored higher than second-borns in curiosity (Tables 1 and 6).

TABLE 3

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: *ENTHUSIASM* AND *AMBITION*

SUBGROUP			*Enthusiasm*				*Ambition*			
Subject		Sibling	AGE DIFFERENCE BETWEEN SIBLINGS IN MONTHS							
Sex	*Ordinal Position*	*Sex*	*7–24*	*25–48*	*49–72*	*Total*	*7–24*	*25–48*	*49–72*	*Total*
			SCORE MEANS				SCORE MEANS			
M	2	M	3.761	4.008	4.515	4.095	3.991	4.230	4.264	4.162
M	2	F	3.698	3.446	3.674	3.606	3.682	3.658	3.720	3.687
F	2	M	4.179	4.531	4.172	4.294	4.173	4.226	4.273	4.224
F	2	F	4.129	3.799	3.824	3.918	3.946	3.759	4.260	3.989
M	1	M	3.779	3.908	4.044	3.910	3.753	3.760	4.129	3.881
M	1	F	3.898	4.354	4.321	4.191	4.043	3.828	4.004	3.958
F	1	M	4.184	4.335	4.654	4.391	4.251	4.346	3.989	4.195
F	1	F	3.918	3.879	4.146	3.981	3.953	3.824	4.001	3.926
M	2	M or F	3.729	3.727	4.095	3.850	3.836	3.944	3.992	3.924
F	2	M or F	4.154	4.165	3.998	4.106	4.060	3.993	4.267	4.106
M	1	M or F	3.838	4.131	4.183	4.051	3.898	3.794	4.067	3.920
F	1	M or F	4.051	4.107	4.400	4.186	4.102	4.085	3.995	4.061
M or F	2	M	3.970	4.269	4.343	4.194	4.082	4.228	4.268	4.193
M or F	2	F	3.913	3.622	3.749	3.762	3.814	3.709	3.990	3.838
M or F	1	M	3.981	4.121	4.349	4.150	4.002	4.053	4.059	4.038
M or F	1	F	3.908	4.117	4.234	4.086	3.998	3.826	4.002	3.942
M	1 or 2	M	3.770	3.958	4.279	4.002	3.872	3.995	4.197	4.021
M	1 or 2	F	3.798	3.900	3.998	3.899	3.863	3.743	3.862	3.823
F	1 or 2	M	4.182	4.433	4.413	4.342	4.212	4.286	4.131	4.210
F	1 or 2	F	4.023	3.839	3.985	3.949	3.949	3.792	4.130	3.957
M	1 or 2	M or F	3.784	3.929	4.139	3.950	3.867	3.869	4.029	3.922
F	1 or 2	M or F	4.103	4.136	4.199	4.146	4.081	4.039	4.131	4.083
M or F	1 or 2	M	3.976	4.195	4.346	4.172	4.042	4.141	4.164	4.115
M or F	1 or 2	F	3.911	3.869	3.992	3.924	3.906	3.768	3.996	3.890
M or F	2	M or F	3.942	3.946	4.046	3.978	3.948	3.968	4.129	4.015
M or F	1	M or F	3.945	4.119	4.291	4.118	4.000	3.940	4.031	3.990
M or F	1 or 2	M or F	3.943	4.032	4.169	4.048	3.974	3.954	4.080	4.003

12. Children whose sib was their junior and opposite in sex were judged to have a greater number of interests and greater enthusiasm than those whose sib was of the same sex. However, in the two traits named above M2M's scored higher than M2F's (Tables 2, 3 and 6).

13. First-born males at the under-two-year spacing received a higher rating in planfulness than second-borns but the trend was not consistent for girls (Tables 5 and 6).

14. The rating given first-born girls on the trait, tendency to dawdle and procrastinate, was higher, on the average, than that given second-born girls. No similar trend was noted for boys (Tables 5 and 6).

Sex-Group Differences (See Items 3, 5, 9, 13 and 14.)

15. Girls were judged more aggressive than boys when the sib gap was less than two years (Tables 1 and 6).

16. Generally girls were rated more enthusiastic than boys but first-born boys with a sister two to four years younger scored higher than the parallel group of girls (Tables 3 and 6).

17. Girls tended to receive a higher rating on responsibleness, tenacity and planfulness than boys. In the first two traits the sex difference was significant at the spacings under four years, in the third, only when the sib age difference was under two years (Tables 4, 5 and 6).

18. Second-born boys were judged to dawdle and procrastinate more than second-born girls but no consistent sex differences are to be noted among first-borns (Tables 5 and 6).

Some Group Characterizations

A. M2M's, as the sib age difference widened, showed more aggressiveness, curiosity, originality, enthusiasm and planfulness and were judged to be possessed of more interests. The trend in the case of the other traits, with the exception of procrastination, was similar but group differences were not large enough to be statistically significant. Generally speaking, then, the wider the sib age difference, the more effective this group type became.

The difference between M2M's and M2F's increased with spacing and was significant at the four-year spacing in the case of aggressiveness, enthusiasm, ambition, tenacity and number of interests, M2M's receiving the higher assessment.

M2M's tended not to differ significantly from parallel groups of M1M's but at the widest spacing scored higher in enthusiasm and at the closest spacing, lower in curiosity.

B. M2F's, in contrast to M2M's, changed little as the age difference between them and their sib expanded and seemed to rank rather low generally. They scored less than M2M's on most traits studied and were rated significantly lower in many at the widest spacing (Item A). When the sib age difference was under two years, the former also scored significantly or

TABLE 4

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: *RESPONSIBLENESS* AND *TENACITY**

SUBGROUP Subject *Sex*	*Ordinal Position*	Sibling *Sex*	*Responsibleness*				*Tenacity**			
			AGE DIFFERENCE BETWEEN SIBLINGS IN MONTHS							
			7–24	*25–48*	*49–72*	*Total*	*7–24*	*25–48*	*49–72*	*Total*
			SCORE MEANS				SCORE MEANS			
M	2	M	3.716	3.773	4.040	3.843	3.736	3.826	4.048	3.870
M	2	F	3.884	3.916	3.685	3.828	3.675	3.932	3.601	3.736
F	2	M	4.569	4.560	3.943	4.357	4.318	4.558	4.020	4.298
F	2	F	4.127	3.741	4.143	4.004	4.413	3.709	4.194	4.105
M	1	M	3.763	3.627	3.954	3.781	3.952	3.761	4.095	3.936
M	1	F	3.786	3.396	3.789	3.657	3.608	3.871	3.913	3.797
F	1	M	3.892	4.195	4.159	4.082	4.169	4.563	4.023	4.252
F	1	F	4.206	4.106	4.117	4.143	3.629	3.996	4.017	3.880
M	2	M or F	3.800	3.844	3.863	3.836	3.706	3.879	3.825	3.803
F	2	M or F	4.348	4.151	4.043	4.180	4.365	4.133	4.107	4.202
M	1	M or F	3.774	3.512	3.872	3.719	3.780	3.816	4.004	3.867
F	1	M or F	4.049	4.150	4.138	4.112	3.899	4.279	4.020	4.066
M or F	2	M	4.143	4.167	3.991	4.100	4.027	4.192	4.034	4.084
M or F	2	F	4.005	3.828	3.914	3.916	4.044	3.820	3.898	3.921
M or F	1	M	3.827	3.911	4.057	3.932	4.060	4.162	4.059	4.094
M or F	1	F	3.996	3.751	3.953	3.900	3.618	3.933	3.965	3.839
M	1 or 2	M	3.739	3.700	3.997	3.812	3.844	3.793	4.072	3.903
M	1 or 2	F	3.835	3.656	3.737	3.743	3.642	3.902	3.757	3.767
F	1 or 2	M	4.231	4.378	4.051	4.220	4.243	4.560	4.021	4.275
F	1 or 2	F	4.166	3.924	4.130	4.073	4.021	3.852	4.105	3.993
M	1 or 2	M or F	3.787	3.678	3.867	3.777	3.743	3.848	3.914	3.835
F	1 or 2	M or F	4.198	4.150	4.090	4.146	4.132	4.206	4.063	4.134
M or F	1 or 2	M	3.985	4.039	4.024	4.016	4.044	4.177	4.046	4.089
M or F	1 or 2	F	4.001	3.790	3.933	3.908	3.831	3.877	3.931	3.880
M or F	2	M or F	4.074	3.998	3.953	4.008	4.035	4.006	3.966	4.002
M or F	1	M or F	3.912	3.831	4.005	3.916	3.839	4.048	4.012	3.966
M or F	1 or 2	M or F	3.993	3.914	3.979	3.962	3.937	4.027	3.989	3.984

near significantly lower than F2F's in originality, number of interests, enthusiasm, planfulness and tenacity as well as higher in dawdling. At the widest spacing M2F's received a lower rating in ambition, responsibleness and tenacity and a higher one in dawdling. The former were usually judged somewhat lower in curiosity, enthusiasm, planfulness and number of interests than M1F's—e.g., at the middle spacing they scored significantly lower in aggressiveness, curiosity and enthusiasm; and at the widest, significantly lower in curiosity, number of interests and enthusiasm. Their picture, then, is that of a somewhat passive, depressed group. Other data we have indicate them to be distinctly sissyish, hostile and not very friendly (10, 11, 12).

C. F2M's showed few significant spacing-group differences. As the sib age difference widened, the ratings in planfulness and dawdling increased. Scores in responsibleness and tenacity were lower at the wide than at the middle spacing. When compared with the F1M's, F2M's rated lower in curiosity and tendency to procrastinate but at the two closer spacings were judged more responsible. They were judged significantly more curious, original, enthusiastic, ambitious, responsible, tenacious and less given to dawdling than the F2F's when the sib age difference was two to four years. They did not differ significantly from the M2M's, except that they scored at the two closer spacings higher in the traits of responsibleness and tenacity, in which girls generally tend to excel; at the middle spacing, higher in enthusiasm; and at the under-two-year spacing, lower in procrastination.

The F2M's, hence, appear to be a rather dynamic group, often scoring higher than any of the other second-borns (this, however, not significantly usually except relative to the M2F's). The total pattern of the data suggests that having an older sib opposite in sex may be more stimulating for the girl than for the boy.

D. F2F's differed little from the F2M's except at the middle spacing (Item C), where they tended to score lower on most of the traits studied. The former differed little also from F1F's. At the wide spacing they were judged to dawdle less; at the close spacing, to be more tenacious than the latter. When the sibs are of the same sex, ordinal position differences seem slight (see M2M's as well).

F2F's showed some spacing differences, being at the two-to-four-year spacing judged significantly less original, tenacious and planful than at the under-two-year spacing and, when the middle spacing is compared with the wide, the former group was rated higher in procrastination and lower in ambition and tenacity. In other words, the F2F's seemed to show a low at the two-to-four-year spacing, while, if anything, the F2M's tended to score relatively high at the same level. We shall comment on this in the discussion.

E. The M1M's, like the M2M's, showed greater aggressiveness the wider the sib age gap. While the spacing-group differences were usually insignificant, the trend in the case of most traits was in a positive direction, as the sib age gap widened—i.e., the score at the widest spacing was usually the highest, though not significantly so. In the main, the M1M's did not differ significantly from the M2M's, but at the close spacing the former were judged the more curious and at the widest spacing, the less enthusiastic.

TABLE 5

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: *PLANFULNESS** AND *TENDENCY TO DAWDLE AND PROCRASTINATE*

SUBGROUP			*Planfulness**				*Tendency to Dawdle and Procrastinate*			
Subject		Sibling	AGE DIFFERENCE BETWEEN SIBLINGS IN MONTHS							
Sex	*Ordinal Position*	*Sex*	*7–24*	*25–48*	*49–72*	*Total*	*7–24*	*25–48*	*49–72*	*Total*
			SCORE MEANS				SCORE MEANS			
M	2	M	3.484	3.769	3.984	3.746	4.099	3.813	3.929	3.947
M	2	F	3.606	3.887	3.919	3.804	4.273	4.126	4.241	4.213
F	2	M	3.816	4.096	4.367	4.093	3.419	3.619	3.913	3.650
F	2	F	4.441	3.972	4.049	4.154	3.831	4.064	3.580	3.825
M	1	M	4.028	3.646	4.104	3.926	3.924	4.166	4.104	4.065
M	1	F	4.283	4.100	3.997	4.127	4.175	3.939	3.894	4.003
F	1	M	4.294	3.989	4.233	4.172	4.314	3.950	4.077	4.114
F	1	F	4.072	3.730	4.138	3.980	4.133	4.038	4.282	4.151
M	2	M or F	3.545	3.828	3.951	3.775	4.186	3.969	4.085	4.080
F	2	M or F	4.128	4.034	4.208	4.123	3.625	3.842	3.747	3.738
M	1	M or F	4.155	3.873	4.050	4.026	4.049	4.052	3.999	4.034
F	1	M or F	4.183	3.859	4.186	4.076	4.223	3.994	4.179	4.132
M or F	2	M	3.650	3.933	4.175	3.919	3.759	3.716	3.921	3.799
M or F	2	F	4.023	3.929	3.984	3.979	4.052	4.095	3.910	4.019
M or F	1	M	4.161	3.818	4.168	4.049	4.119	4.058	4.091	4.089
M or F	1	F	4.178	3.915	4.068	4.053	4.154	3.988	4.088	4.077
M	1 or 2	M	3.756	3.708	4.044	3.836	4.011	3.989	4.017	4.006
M	1 or 2	F	3.944	3.994	3.958	3.965	4.224	4.032	4.067	4.108
F	1 or 2	M	4.055	4.042	4.300	4.132	3.867	3.784	3.995	3.882
F	1 or 2	F	4.257	3.851	4.094	4.067	3.982	4.051	3.931	3.988
M	1 or 2	M or F	3.850	3.851	4.001	3.900	4.118	4.011	4.042	4.057
F	1 or 2	M or F	4.156	3.947	4.197	4.100	3.924	3.918	3.963	3.935
M or F	1 or 2	M	3.905	3.875	4.172	3.984	3.939	3.887	4.006	3.944
M or F	1 or 2	F	4.101	3.922	4.026	4.016	4.103	4.042	3.999	4.048
M or F	2	M or F	3.837	3.931	4.080	3.949	3.905	3.905	3.916	3.909
M or F	1	M or F	4.169	3.866	4.118	4.051	4.136	4.023	4.089	4.083
M or F	1 or 2	M or F	4.003	3.899	4.099	4.000	4.021	3.964	4.002	3.996

M1M's seem a rather neutral group. In general, they were judged somewhat lower than M1F's but not significantly. They scored significantly less than the latter, however, at the middle spacing on aggressiveness, curiosity, enthusiasm and planfulness.

F. M1F's increased significantly from the close to the middle spacing in aggressiveness, curiosity and number of interests. At the under-two-year sib-age difference the group seemed rather passive. However, on practically all of the positive traits studied, M1F's scored above (not always significantly above, however) the M2F's at the spacings wider than two years, but at the middle spacing the former were rated less in responsibleness. When child and sib were less than two years apart in age, M1F's scored higher in planfulness. The M1F type seemed most on its toes when the sib age gap was two to four years—i.e., here they scored higher than M1M's on aggressiveness, curiosity, enthusiasm and planfulness and higher than F1F's on aggressiveness, curiosity and enthusiasm. The former were rated lower, however, on responsibleness than the latter. The outstanding pattern detail in the case of M1F's is, in our opinion, the change in a positive direction from the close to the middle spacing.

G. F1M's, generally, tended to rate higher than the other groups on most of the traits but the group differences were not always significant. The evaluation of F1M's curiosity was significantly higher than that of the F2M's at all spacings. The former was gauged at the wide spacing the more enthusiastic; at the close spacing, as more given to procrastination and as less responsible. F1M's scored significantly higher than the F1F's at some level on practically all of the traits and rather consistently higher on all of the traits except responsibleness and procrastination. From these observations we suspect that a male sib sparks the girl's drive. The girls with a younger brother received significantly higher ratings in enthusiasm than the boys with a brother their junior. At the close spacing the former group was judged more aggressive and ambitious and at the middle, more ambitious, tenacious and responsible. The group differences at the wide spacing were usually insignificant. The F1M's seem the most dynamic of our groups as far as the traits studied are concerned.

H. The F1F's, in contrast to the F1M's, were a rather average or steady group, changing insignificantly with spacing. They differed little from the F2F's, except at the close spacing, where they received a lower rating on tenaciousness and at the over-four-year spacing, where they were judged to procrastinate the more. The F1F's seemed less on their toes than the F1M's, scoring lower generally, except in procrastination and responsibleness. At one spacing or another F1F's scored significantly lower on aggressiveness, curiosity, enthusiasm, ambition and tenacity. When compared with the M1F's, the former were assessed the more responsible. F1F's were rated more aggressive at the close spacing and at the middle, less curious and enthusiastic.

TABLE 6

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	AGGRESSIVENESS			CURIOSITY			ORIGINALITY			NUMBER OF INTERESTS			ENTHUSIASM		
		Sum of Squares	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F
Total	383	330.80			309.15			332.19			312.58			313.56		
Between groups	23	22.91	1.00		24.88	1.08		10.20	.44		18.48	.80		34.25	1.49	
Within groups	360	307.89	.86		284.27	.79		321.99	.89		294.10	.82		279.31	.78	
Between replications	15	25.01	1.67		17.06	1.14		13.18	.88		20.72	1.38		12.24	.82	
Residual	345	282.88	.82		267.21	.77		308.81	.90		273.38	.79		267.07	.77	
Between Variables:																
Between sex groups	1	1.61	1.61		.02	.02		.96	.96		.31	.31		3.47	3.47	4.48
Between sibling's-sex groups	1	2.46	2.46		.81	.81		1.32	1.32		1.35	1.35		6.49	6.49	8.39
Between ordinal-position groups	1	.35	.35		7.73	7.73	9.98	.04	.04		.77	.77		1.59	1.59	
Between spacing groups	2	3.14	1.57		3.99	2.00		.18	.09		5.10	2.55	3.22	3.40	1.70	
First-order Interactions:																
Subject's sex—sibling's sex	1	.00	.00		2.57	2.57		.33	.33		.01	.01		2.16	2.16	
Subject's sex—ordinal position	1	.18	.18		.02	.02		.70	.70		.00	.00		.41	.41	
Subject's sex—spacing	2	3.40	1.70		.67	.33		.78	.39		1.96	.98		1.07	.53	
Sibling's sex—ordinal position	1	.06	.06		.19	.19		.02	.02		.32	.32		2.86	2.86	
Sibling's sex—spacing	2	3.09	1.55		.05	.03		2.44	1.22		.61	.31		1.82	.91	
Ordinal position—spacing	2	.19	.09		.18	.09		.21	.11		.62	.31		.94	.47	
Second-order Interactions:																
S's sex—sib's sex—ordinal position	1	2.54	2.54		1.86	1.86		1.04	1.04		3.53	3.53	4.45	4.09	4.09	5.28
S's sex—sib's sex—spacing	2	.65	.33		2.37	1.19		.35	.18		.62	.31		.87	.44	
S's sex—ordinal position—spacing	2	3.72	1.86		.72	.36		.16	.08		.35	.18		2.69	1.35	
Sib's sex—ordinal pos.—spacing	2	1.16	.58		2.85	1.43		.64	.32		2.43	1.22		1.60	.80	
Third-order Interaction:																
S's sex—sib's sex—ord. pos.—spac'g	2	.37	.18		.84	.42		1.02	.51		.49	.25		.77	.38	

TABLE 6 (*continued*)

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

		AMBITION			RESPONSIBLENESS			TENACITY			PLANFULNESS			TENDENCY TO DAWDLE AND PROCRASTINATE		
Source of Variance	*df*	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F	*Sum of Squares*	*Estimate of Variance*	F
Total	383	333.93			338.84			349.52			347.36			336.13		
Between groups	23	17.37	.76		29.24	1.27		28.86	1.25		21.23	.92		19.23	.84	
Within groups	360	316.56	.88		309.60	.86		320.66	.89		326.13	.91		316.90	.88	
Between replications	15	20.03	1.34		17.00	1.13		17.10	1.14		29.13	1.94		12.30	.82	
Residual	345	296.53	.86		292.60	.85		303.56	.88		297.00	.86		304.60	.88	
Between Variables:																
Between sex groups	1	2.26	2.26		13.62	13.62	16.06	8.24	8.24	9.36	3.82	3.82	4.44	1.29	1.29	
Between sibling's-sex groups	1	4.83	4.83	5.62	1.12	1.12		4.45	4.45	5.06	.13	.13		1.05	1.05	
Between ordinal-position groups	1	.06	.06		.82	.82		.17	.17		1.08	1.08		2.92	2.92	
Between spacing groups	2	1.15	.58		.45	.22		.43	.21		2.41	1.20		.21	.11	
First-order Interactions:																
Subject's sex—sibling's sex	1	.12	.12		.09	.09		.60	.60		.90	.90		.01	.01	
Subject's sex—ordinal position	1	.08	.08		.10	.10		1.07	1.07		2.14	2.14		4.92	4.92	5.57
Subject's sex—spacing	2	.20	.10		1.25	.62		1.03	.52		.70	.35		.29	.14	
Sibling's sex—ordinal position	1	1.64	1.64		.56	.56		.25	.25		.05	.05		1.29	1.29	
Sibling's sex—spacing	2	1.03	.51		1.14	.57		.66	.33		1.89	.95		.59	.30	
Ordinal position—spacing	2	.36	.18		1.00	.50		1.14	.57		2.59	1.30		.20	.10	
Second-order Interactions:																
S's sex—sib's sex—ordinal position	1	2.30	2.30		1.85	1.85		.23	.23		.94	.94		.28	.28	
S's sex—sib's sex—spacing	2	1.97	.98		2.09	1.05		6.49	3.24	3.69	1.02	.51		.71	.36	
S's sex—ordinal position—spacing	2	1.18	.59		1.71	.86		2.08	1.04		.83	.42		1.64	.82	
Sib's sex—ordinal pos.—spacing	2	.02	.01		.45	.22		1.54	.77		1.15	.58		.83	.42	
Third-order Interaction:																
S's sex—sib's sex—ord. pos.—spac'g	2	.17	.08		3.01	1.50		.49	.25		1.57	.78		3.00	1.50	

Discussion of Results

Spacing- and Sib's-Sex-Group Differences

It seems reasonable to us that number of interests should increase as the age difference between the sibs increases (Item 1). The less alike in age the child and his sib are the more the former is likely to be thrown for companionship with children not in his home. This should extend his interests. Furthermore, in the case of the second-born, the effort of the child to keep up with his older sib (whose interests should be more numerous the older he is) should also expand the former's interests as the sib age difference widens.

That tenacity should be shown to a greater degree by the rivalrous, jealous, competitive aggressive girl with a younger brother (10, 11, 12) than by the girl with a younger sister (Item 2), one might also have predicted. The drive of the F1M's, we think, is sparked by sex rivalry; and the confidence which comes from being the abler member of the sib pair probably also makes for high persistence in effort as well as dominance on F1M's part. Since at the four-to-six-year spacing, when the sib of our five-and-six-year-old subjects would often be a mere infant and not a competitor either in the matter of skills or possessions, nor its sex conspicuous, we might expect, what we found, that the tenacity difference of F1M's and F1F's was insignificant. At the spacings under four years, however, when sib interests overlap considerably, yet the treatment differential between the siblings must be great, and this especially if they differ in sex, sheer persistence may be more rewarding in adjustment than it is when the spacing is wide and contact with the sib very limited. F1M's, who are very jealous of the brother (12)—possibly with good reason because boys tend to be indulged more by their mothers (17)—and expansive in their dealing with the problem, might be expected to be more tenacious than the F1F group.

That the F2M's and F2F's differed only at the two-to-four-year spacing (Item 7) may be due to the possibility the members of the two groups associated to a different degree with their sibs at this level. The phenomenon of sex distance, which is particularly high among grade school youngsters (9), may isolate F2M much from her two-to-four-year-older sib and his associates, while F2F will be functioning chiefly as a much out-classed hanger-on in her sister's play groups. F2M(25-48)'s, having more playmates of their own choosing, are likely to have their confidence, drive and rivalry stimulated by the latter, while F2F(25-48)'s, being chiefly hangers-on still in the sib's groups, feel too inferior to be very tenacious or dominant.

In the case of the second-born boy we seem not to have a similar picture. The forces which make the boy with a much older brother more tenacious than the one with a much older sister (Item 2) we have in a sense discussed elsewhere (12) but let us comment briefly here. (See Items A and B; also Items 2, 4, 6, 8, 10 and 12.) The boy with an older brother identifies with him and wants to keep up with the latter's standards. His sib reinforces his sex identification and reduces conflict (12). But when the age difference between the child and his older brother is over four years, their play groups overlap little. The brother is out of the home much of the time. Hence M2M does not have the experience of being constantly outclassed and has playmates of his own age, sex and choosing, who bolster his morale, and with whom he can compete as an equal. His mother, furthermore, probably protects him much against his much older brother and his father may also, there being no cross-sex alliances in this case. Hence M2M's experience of being outstripped is not too frequent. This constellation of influences should strengthen M2M(49-72)'s drive, self-confidence and initiative.

M2F, on the other hand, has much sex identification conflict, we believe. The models he has most contact with (his sib and mother) are females. If Freudian theory is correct, his father probably favors the sister. The sib is likely, too, to bar her four-to-six-year younger brother from her play groups, yet she is at home much and is probably often put in charge of the younger brother with whom she is very rivalrous, dominating and possibly even protective (see F1M and also 11, 12). His feeling of inferiority, the greater passivity of the female model he has, his father's likely favoring of the sib, his more frequent association with his sib who beats him down and his mother's indulgence of him because he is a male, the youngest, and different in sex from the older sib, we suspect, curb the drive and tenacity or dominance of M2F. That he is encouraged in a more passive pattern seems likely. He may have compensation for his feelings of inferiority in the indulgence and babying he receives from his mother and others. Hence we would expect, when the sib age gap is wide, M2F would be less tenacious and less dynamic generally than M2M; also that the difference between the two groups would increase as the sib age gap widens. It is to be noted that the boy with a much older brother, compared with one with a much older sister, was judged the more aggressive, enthusiastic, ambitious, tenacious and possessed of the more interests (Items 2, 4, 9 and 12).

It is interesting that M2M's differed insignificantly from M1M's, except that at the widest spacing the boy with the much older brother was judged to be more enthusiastic.

The variable, sex of sib, in relation to sex of child, seems an important conditioner of many of the group differences in personality which we have observed (Items 6 to 10 and A to H). The girl with a younger brother, for instance, apparently stimulated by sex rivalry, presents a picture of greater dynamicness than the girl with a younger sister. At some spacing in the case of practically all of the traits, except responsibleness and procrastination, the assessment the F1M's received was significantly higher than that the F1F's received.

That F1M is so confidently rivalrous and expansive (11, 12) we think arises out of her identification with her mother—possibly also out of attentions from her father—and the fact that, as the older sib, she is the abler. She believes her sex to be the better (1, 20); but, challenged by her brother, she has more need to prove herself "best" than has F1F. The former's younger brother (see M2F), whatever his age, is probably sufficiently favored and indulged by the mother (17 and also Item B) so that she is kept ever alert. That the age difference between F1M and her brother, and the treatment difference this implies, does not alter in a very consistent way the work habit differences between her and the F1F's, suggests it is the sheer fact of the sib's difference in sex and the probable attitude of the mother, etc., associated with this that is important rather than characteristics of the sib other than his sex.

It impresses us that at the two-to-four-year spacing sib's-sex-group differences seem particularly marked (Items 2, 7, 9, 10 and 16). M1F(25-48)'s, for instance, received a higher rating than M1M(25-48)'s in aggressiveness, curiosity, enthusiasm, and planfulness (Items E and F); F1M's exceeded F1F's in curiosity and enthusiasm (Items G and H); F2M's scored higher than F2F's in curiosity, originality, enthusiasm, ambition, responsibleness, tenacity, and lower in procrastination (Items C and D). The exception to the trend of greater apparent drive in the child at the middle spacing whose sib is opposite in sex is the boy with an older sister, who either differs insignificantly from the boy with an older brother or in the depressed direction. We have already discussed the case of M2F and attributed his passivity probably to the excessive babying he receives from his mother, the dominance of his jealous older sister and his sex identification conflict.

The forces which motivate the first-born whose sib is two to four years younger, we think, may assume a pattern such as the following: Since our subjects were five- and six-year-olds, the sib would be at that adorable one-to-four-year age, when it receives much attention and appreciation from the family and others. Yet it is not too satisfying a companion for the older child because of its immaturity. It gets around on its own and in its clumsy way may interfere much with the activities and possessions of the latter.

Some would want to put more emphasis on the *time* in the older child's life of the arrival of the sib, or when the latter becomes a competitor, and this in relation to the possibility of serious interference with the former's dependency needs at a critical stage in its development. Stendler (19) has the belief, for instance, that the child from nine months to three years is particularly susceptible to trauma with persisting effects. Evidence is accumulating that the effects of trauma occurring in an individual with very restricted experience tend to be very durable.

If the sib is opposite in sex, the treatment differential will be greater than if the sib is of the same sex as the child and this differential will probably be great at the middle spacing and perhaps seem to the child less justified than would be the case were the sib much older or younger. The sex of a very young sib will not be an important determinant of differences in sib treatment, and if the sib is very near the older child in age, a low difference in treatment of the two children probably obtains as well, parents tending to interfere less in the sib altercations if there is a chance of the children settling matters by themselves. Parents will need to interfere most at the middle spacing and in doing this will doubtless reveal their sex bias, or the child, looking for reasons for a difference in treatment, will see this difference as due to his sex difference. The high jealousy of M1F(25-48)'s (11), we believe, may lie back of his higher score in curiosity, aggressiveness and planfulness (Item 7). We also have the hunch that at the middle spacing the sex difference between the sibs makes the most difference in the degree to which the child has associates of his own sex, age and selection, a condition which we suspect makes for greater expansiveness and a more dynamic attitude on the part of the subject. If the sib is an infant or very much older, the great maturity difference between the sibs throws the child with associates of his own, no matter what the sex of his sib. On the other hand, should the sib be very near in age, then he will tend to have the same playmates—i.e., a sex difference will probably not be an important factor in assuring the child associates of his own. However, when the age difference is two to four years, the maturity difference between the sibs, in addition to a sex difference, may separate some the sibs opposite in sex, whereas the maturity difference alone would not. In other words, we think it takes a larger maturity difference to separate the sibs substantially, if it operates alone.

A child two to four years younger than his sib and of the same sex is likely to have the role of a hanger-on in groups chosen chiefly by the latter, whereas, if the child differs in sex from his sib, he is more likely to have playmates of his own selection, a condition which we think builds his confidence, bolsters his morale, makes him stand on his own feet and encourages him to be socially more expansive. As we have said, the fact that M2F does not seem to fit this pattern may be due to his being indulged and protected excessively by his mother—possibly even by his older sister as well—and kept apart. There is some evidence the mother babies him, encourages dependence and creates a state of mind unfavorable to interest in peers (12). M2M at the widest spacing, when we would expect him to have a coterie of friends of his own and to be encouraged in the masculine pattern, is, in contrast, buoyant. At the middle spacing when M2M is still chiefly a hanger-on, he tends to be rather ineffective, resembling F2F, who at this spacing would be in a similar role. M2M's and F2F's at the two-to-four-year sib-age-difference level were, one or the other, rated significantly lower in curiosity, enthusiasm, ambition, tenacity, than the corresponding group at the widest spacing.

If the sib is younger, the sex distance factor (9), we believe, will operate less potently because both children will be at home most of the time and playmate groups will overlap. Should the sib be an infant, however, it will not be much of a participant in the older child's play and the latter's playmates will be chiefly of his own choosing. But we feel sure the sex distance factor does operate some at the middle spacing. The boy and his friends may be thought too rough by the one-to-four-year sister, for instance, and she may seek associates of her own sex. If the sibs' age disparity is less than two years, their play groups will doubtless overlap much or be virtually identical.

We suspect, then, the greater aggressiveness, curiosity, enthusiasm and planfulness of M1F(25-48) than M1M(25-48) (Items 7, 12, E and F) is related in part to his having a coterie of friends of his own and his rivalrousness due to the adoration the sib receives, which is unique to the latter's age.

The greater drive or tension generally of F1M than F1F (Items G and H and 2, 6, 7, 8, 9 and 12) probably demands some comment. It is our belief the first-born girl identifies strongly with her mother. This should give the former considerable confidence about her sex classification but she is at the age when typically girls become increasingly interested in their fathers and in sex roles. Hence it is our hunch the girl is less likely to question her own worth if her sib is of her own sex. Should the sib be different, the girl may have to prove to herself and others she is the better. If the sib is male, she may sense vaguely the mother's favoring of the brother (Freudian doctrine maintains the latter is the usual tendency) and even the premium placed by the culture on the male. Against doubt the girl is likely to react vigorously and, in the main, by reaching out. F1M tends to be more aggressive, curious, enthusiastic, tenacious than F1F, for instance. The activeness, hostility and vigor of the brother may be another determinant of F1M's reaction pattern but this could scarcely be very potent at the wider spacings.

F2M reacts somewhat less vigorously than F1M, we suspect, because she is outclassed by her older brother from the first and she has not experienced being superseded by her sib. Many would insist F1M's drive is somewhat anxiety motivated, because of the displacement experience and possibly anxiety contamination from her inexperienced parents whose first child she was. Hence F2M would be viewed as more relaxed than F1M rather than less dynamic. It is possible F1M may have been given more responsibility and been stimulated more than F2M, the latter remaining the youngest in the family, since we dealt only with two-children families. What interests us is that if the latter forces listed are very potent, they should produce ordinal-position differences in all the groups. These differences seem to show, however, only in the case of the F1M's and M1F's—a fact which suggests that considerable weight should be given to the idea that the first born, who is particularly status-sensitive, has his drive fed very significantly by sex rivalry.

We have noted that those with a male sib were judged more ambitious, on the average, than those with a female sib (Item 9). If we look closely at the tables, however, we see that M1M presents an exception to this trend and that only the girls with a brother and M2M at the wider spacings fit the pattern. Hence we suspect it is not the mere maleness of the sib that is the potent mechanism in these results. The girl with a brother, as has been mentioned earlier, it is our belief, is jealous and is spurred by sex rivalry. This could make her more ambitious than the girl with a sister, as well as more enthusiastic and aggressive. That M2M is more ambitious than M2F at the wider spacings we suspect results from forces quite other than direct sex rivalry. As has been explained, M2F probably suffers seriously from a sex identification problem but seems to deal with it rather passively, getting some compensation for this possibly from the indulgence of his mother and even of his older sister. Instead of having his drive increased by sex rivalry, it is our hunch, he is depressed and withdraws. In this latter he may be somewhat encouraged by his indulgent mother. He may also be too indulged to be able to tolerate the buffets of his peers. He ranks low in ambition, enthusiasm and initiative. Back of the trend in all three traits we think we see the evidence of the same set of influences (see also Items 4, 6, 8, 10 and 12).

Ordinal-Position Group Differences

It is interesting first-borns tended to be judged more curious than second-borns (Item 4). This may be due to the fact the former have associated more with adults, who have aroused their interests in many things and encouraged their questions. If this were the mechanism, we should expect curiosity to increase with spacing both for first- and second-borns. This is, to be sure, the trend but the spacing-group differences are not significant.

It is also possible that curiosity is greatest when the child is on the *qui vive*. The first-born, because he has been displaced, is the more concerned about his rights and status. Our evidence that the older member of the pair tends to be the more competitive and insistent on his rights (12) gives the hypothesis some support. It is relevant, too, that among the first-borns, those in opposite-sex pairs, where the sib's threat to status is greatest, showed the most curiosity (Item 7).

First-borns at the close spacing were rated more planful, too, than second-borns. The former, when the sib follows close on his heels, does not long have the experience of being the sole object of parental concern. The second-born, in our two-child families at least, remained, on the other hand, the youngest and hence probably longer had the type of service and protection we are wont to call "babying." Responsibility is put on to the shoulders of the first-born sooner, we think. All this should encourage him in being planful. Since his competitor is formidable, too, if the spacing is close, he can maintain the top position only if he plans carefully and looks ahead. Planfulness may hence be very rewarding. We have the hunch, too, that the first-born tends to be contaminated by his parents' caution and uncertainty in their new role as parents. Fearfulness motivates planning, looking ahead, and makes for restraint and curbing of impulsiveness. That the first-born was judged the more planful only when his sib differed little in age we suspect may result from influences at the wider sib spacings which counter to some extent those just described as favoring the older child. The first-born, for instance, who was long an only child might have had the initiative in planning kept from him for a sufficient time to weaken his inclination.

It did not surprise us when scanning ordinal-position differences that first-born girls were thought to dawdle more than second-born. Procrastination and dawdling may be one of the less aggressive ways of luring to the self the longed-for service and support. The first-born girl among our subjects has functioned as an only child for from nine months to six years. She has had her parents' attention undiluted and has probably come to be more dependent on it than the second-born who has always had to share attention and service. But what we find puzzling, if our hunch is correct, is that first-born boys were not rated higher in procrastination and dawdling than second-borns. Since girls probably are indulged less (17), and when they are the older sib are given much responsibility relatively, they may have more need to use dawdling as an attention- and service-getting device than the boy. Girls also would tend to favor the more passive kinds of adjustment of which dawdling is one. We note, however, when we scan the data in detail, that it is M2F, who we think is pampered, passive, and possibly inferior-feeling, who ranks highest in dawdling. He doubtless uses dawdling as a device to keep adults at his service, and his mother, being indulgent, probably rewards his efforts. His weight doubtless causes the second-born boys to average as high as the generally more susceptible first-borns.

Sex-Group Differences

The sex differences our data show tend to be in line with those reported in the literature. The girl was judged by her teachers as more responsible, planful and persistent. This finding is congruent with the usual finding that the girl is more obedient, conscientious, and conforming than is the boy. (See References 2, 3, 6, 11, 16 and 20 as examples.) What the relative weights of cultural and biological forces are in determining the difference is, of course, a question.

When the sib was less than two years younger, first-born girls were judged to have more initiative or be more dominant than boys. Since girls are more conforming than boys, the former are probably given more responsibility (17) and are hence also more comfortable at school than is the boy. This may cause the girl to take more initiative, especially in school activities. It was school behavior, we should remind the reader, on which the teachers' ratings were chiefly based. It is also possible that the girl's identification with her mother, and through generalization, with her teacher, that causes her to show more initiative at school. But why, then, would the sex difference obtain only at the close spacing? We think it possible the arrival of a sib close on his heels may traumatize the boy more than the girl since he cannot identify completely with his mother. Threatened more, he may adapt by direct vigorous clinging to his mother whereas the girl may reach out more. It is interesting in connection with this hypothesis that the boys who we think have the least in the way of a sex identification problem—namely, the MM's at the widest spacing—were rated particularly high in initiative. The hunch that the girl at the close spacing is more self-confident than the boy is given support by the teachers' ratings on this trait. Significant also are Tuddenham's (20) and Ausubel's (1) findings which suggest that young girls tend to have a better opinion of themselves than do young boys.

Second-born boys were thought to dawdle more than girls. This may be the result of the mother's greater indulgence of the male and the males' lesser comfort at school. Just why, then, first-born boys were not judged to procrastinate more than the girls is not clear. Possibly the sex differential in indulgence is greater among second- than first-borns.

As a final comment let us point to the sib's-sex variable as a very important one that, in the main, has been relatively neglected in the experimental, though not in the clinical literature. We have been impressed also by the fact that most of the significant relations in our data appear in the interactions. This would justify warning against sweeping generalizations about the effects of birth-order, sib's sex, sex of child or the age difference between sibs.

REFERENCES

1. Ausubel, D. P., *et al.* Perceived parent attitudes as determinants of children's ego structure. *Child Develpm.*, 1954, 25, 173-183.
2. Berne, E. V. C. An experimental investigation of social behavior patterns in young children. *Univer. Iowa Stud. Child Welf.*, 1930, 4, No. 3.
3. Caille, R. K. Resistant behavior of preschool children. *Child Develpm. Monogr.*, 1933, No. 11.

4. Conrad, H. S. *The California Behavior Inventory for Nursery School Children.* Berkeley, California: Univer. of California Press, 1933.

5. Conrad, H. S. A statistical study of ratings on the California Behavior Inventory for Nursery School Children. *Genet. Psychol. Monogr.,* 1934, 16, 1-79.

6. Dean, A. The relation of ordinal position to personality in young children. Unpublished master's thesis, State Univer. of Iowa, 1947.

7. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.

8. Goodenough, Florence, & Anderson, J. E. *Experimental child study.* New York: Century, 1931.

9. Koch, Helen L. A study of some factors conditioning the social distance between the sexes. *J. soc. Psychol.,* 1944, 20, 79-107.

10. Koch, Helen L. The relation of "primary mental abilities" in five- and six-year-olds to sex of child and characteristics of his sibling. *Child Develpm.,* 1954, 25, 209-223.

11. Koch, Helen L. The relation of certain family constellation characteristics and the attitudes of children towards adults. *Child Develpm.,* 1955, 26, 13-40.

12. Koch, Helen L. Attitudes of children toward their peers as related to certain characteristics of their sibling. *Psychol. Monogr.,* 1956, 70, No. 19 (Whole No. 426).

13. Lasko, Joan K. Parent behavior toward first and second children. *Genet. Psychol. Monogr.,* 1954, 49, 97-137.

14. Mayer, A. Chicago: Socio-economic groups of the population by census tracts. The Chicago Community Inventory: Univer. of Chicago, 1940.

15. Reynolds, M. M. Negativism of preschool children. *Teach. Coll. Contr. Educ.,* 1935, No. 288.

16. Richards, T. W., & Powell, Marjorie. The Fels Child Behavior Scales. *Genet. Psychol. Monogr.,* 1941, 24, 259-311.

17. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, P. S. Some child rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.,* 1953, 47, 135-236.

18. Sewall, Mabel. Two studies in sibling rivalry. I. Some causes of jealousy in young children. *Smith Coll. Stud. Soc. Work,* 1930, 1, 6-22.

19. Stendler, Celia B. Possible causes of overdependency in young children. *Child Develpm.,* 1954, 25, 125-146.

20. Tuddenham, R. D. Studies in reputation: I. Sex and grade differences in school children's evaluations of their peers. *Psychol. Monogr.,* 1952, 66, No. 1 (Whole No. 333).

THE EFFECT OF VERBAL TRAINING ON INITIAL GENERALIZATION TENDENCIES

WINIFRED O. SHEPARD
Iowa Child Welfare Research Station

It has often been observed that *Ss* who learn a given response to a stimulus tend to generalize it to other stimuli which are physically similar to it. The amount of generalization decreases as the similarity between training and test stimuli decreases. This has been called primary generalization.

There are, however, other bases for generalization. Birge (1), Jeffrey (4), Murdock (5), Eisman (3) and Shipley (6) have all demonstrated that learning a common response for two or more stimuli increases the tendency to generalize other responses among them. This is known as mediated or secondary generalization.

In these studies, however, not much attention has been paid to the nature of the stimuli involved and so a question is raised as to whether or not there is a relationship between the similarity of the stimuli and the effectiveness of a mediating response. Concretely one might ask, "When *Ss* are taught a common response to two highly dissimilar stimuli, is the tendency to generalize other responses between these stimuli raised to the same level as when *Ss* are taught a common response to two less dissimilar stimuli?"

The present study is a beginning attempt to answer this question. It is concerned with getting an indication as to whether or not a relationship exists between stimulus similarity and the effectiveness of a mediating response. To do this, a training stimulus and a series of test stimuli of increasing dissimilarity from it were used. An attempt was made to determine first the initial or primary generalization gradient and then to see how the gradient would be affected by teaching *Ss* a common name for the training stimulus and each of the test stimuli. It was predicted that such training would increase the amount of generalization at all test points, but no prediction was made as to the amount of increase that would be observed at each point.

SUBJECTS

Fifty-nine children from the Iowa Preschool laboratories and the Parent's Playschool in Iowa City were used as subjects. All were between the ages of 3½ and 6 years at the time of participation.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

APPARATUS

The apparatus was a rectangular black box 14½ in. wide, 17½ in. high and 10½ in. deep. A circular flashed opal glass opening 4 in. in diameter was located in the upper center front of the box. Inside the box, behind the aperture, was a set of wooden compartments each containing a 75-watt projector lamp. It was possible to insert colored glass filters into grooves in front of these bulbs and thus project a colored light onto the glass in the aperture. A series of such lights could be presented manually by operating a push button or automatically through the activation of a stepping relay by two Hunter Timers.

There was also a small push button set in a wooden block which was placed on the table between *S* and the apparatus. Pressure on this button constituted the response measure.

STIMULI

Five 2 by 2 in. colored glass filters manufactured by the Corning Glass works were used as stimuli. They were a red, No. 2418, an orange, No. 3480, a yellow, No. 3307 and a blue, No. 5030. Six staff members and graduate students at the Child Welfare Research Station judged that the stimuli as listed above constituted a series of decreasing similarity from the red.

PROCEDURE

The *Ss* were randomly divided into three experimental groups of 18 each and one control group of five. Each experimental group was tested twice for its tendency to generalize the button pushing response from the training stimulus to one of the test stimuli. The first test was given before *S* learned a common name for the training and test stimuli. The second test was given afterwards. The control group was also given two generalization tests, using the third test point, but members of this group had no intervening verbal training.

Each *S* was brought into a darkened experimental room and seated at a small table. Directly in front of him was the push button. About 18 in. behind that was the main apparatus with the aperture slightly above eye level.

Stage 1 for all *Ss* consisted of 24 random presentations of the red and blue stimuli. *Ss* were taught to push the button when the red light came on and not to push it when the blue light came on. *E* gave verbal reinforcement, "Right," or punishment, "Wrong," after each response throughout this stage.

Stage 2 for all *Ss* was an initial generalization test. Three presentations of the appropriate test stimulus for each group were randomly interspersed among the red and blue stimuli in a series of 12 presentations. No reinforcement for any response was given during this series. During both Stages 1 and 2 the lights were automatically presented with a 4-second intra- and inter-presentation interval.

Stage 3 was given to all experimental groups, but omitted for the control group. *Ss* were given 24 presentations during which the red, blue and test stimuli were randomly interspersed. They were taught the name "Mo" for the red and the test stimuli and the name "Lee" for the blue stimulus. Then there was a rest period during which *Ss* were allowed to choose from among a group of dime store items the "prize" they wished to win when the session was over. After the rest, they were again given the three light series and asked for the name each time. Control *Ss* were merely given the rest period and allowed to choose their prizes.

Stage 4 for the experimental groups consisted of 10 random presentations of the red and blue lights. *Ss* were required to say the name of each light as soon as it came on and then to push the button for red and not to push for blue. As in Stage 1, verbal reinforcement was given throughout. Control group subjects were given these 10 presentations but were not, of course, required to say any names.

Stage 5, the final generalization test, was given to all subjects and duplicated Stage 2, except that experimental subjects were required to say the name of each light as it came on. During Stages 4 and 5 the lights were automatically presented. They were on for six seconds and off for four. The increase in the amount of time the lights were on was necessary to allow the experimental groups time to verbalize the names.

RESULTS AND DISCUSSION

The results are given in Table 1 and are graphically represented in Figure 1. Table 1 shows the proportion of pushing responses made to each stimulus on both the initial and final generalization tests. These figures represent the number of times the pushing response was made to each stimulus divided by the number of times that stimulus was presented. Figure 1 shows the mean number of generalized pushing responses made to each of the test points both before and after verbal training. It was necessary to use proportions in Table 1 since the number of times the red and blue stimuli were presented differed from the number of times the test stimuli were presented. Figure 1, however, deals only with the test points, each of which was presented an equal number of times; thus means are used in presenting the data. It is immediately apparent from both the table and the figure that, in terms of observed results, the prediction was confirmed. At every test point there was more generalization after than

TABLE I

PROPORTION OF PUSHING RESPONSES MADE TO EACH STIMULUS BY ALL GROUPS ON INITIAL AND FINAL GENERALIZATION TESTS

	Red	*Blue*	*Test Point 1*	*Test Point 2*	*Test Point 3*
Initial Test					
Experimental Groups	.98	.01	.80	.24	.22
Control Group	.96	.00	..	..	.33
Final Test					
Experimental Groups	.97	.03	.91	.44	.65
Control Group	1.00	.00	..	..	.33

before verbal training. The control group which received no verbal training showed no change at all.

In view of the extreme skewness of the data at each point, Dixon and Mood's nonparametric Sign Test (2) was used to test for statistical significance. For each experimental group the amount of generalization before and after verbal training was compared. *S*s tested at point 1 showed no significant increase. *S*s tested at points 2 and 3 did show significant in-

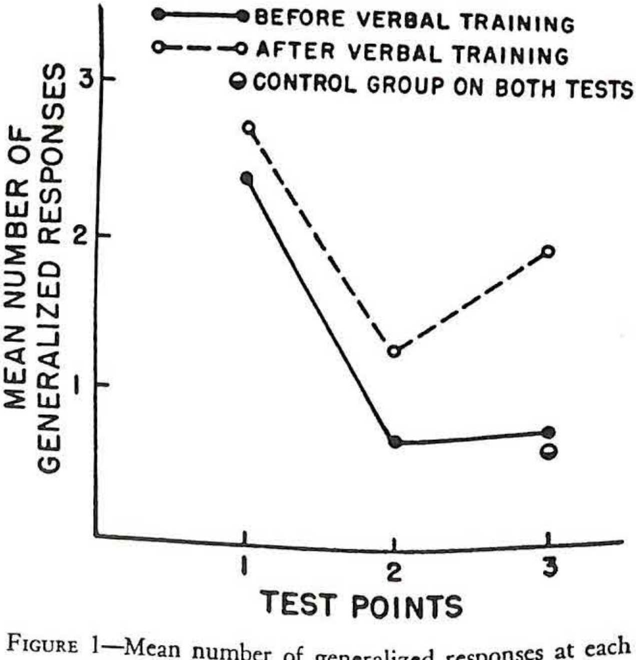


FIGURE 1—Mean number of generalized responses at each test point.

creases with the respective probabilities of occurrence being 5 per cent and .05 per cent. It seems likely that the failure to find a significant increase at point 1 can be attributed to a ceiling effect. *Ss* tested at this point had such strong generalization tendencies initially that there was little opportunity to demonstrate an increase. In view of the small *N* in the control group, no test of significance was made. It is noteworthy, however, that not one of the *Ss* in this group changed at all from the first to the second test indicating that mere continued exposure to a set of stimuli will not, per se, result in increased generalization.

The notion of mediated generalization has been commonly used to explain generalization increases of the type found in this study. According to this view, the mediating response, in this case the name "Mo," itself produces a stimulus which occurs in conjunction with each of the external stimuli to which it has been conditioned. The presence of this common response-produced cue in two (or more) stimulus complexes is said to make these complexes more similar to each other than are the external stimuli alone and hence there is more tendency to generalize among them. It follows then that the optimum situation for demonstrating mediated generalization is one in which *Ss* are forced to make the mediating response overtly during the generalization test, thus maximizing the probability that the response-produced cue will be present. These conditions were met in this study. Nevertheless, about half the *Ss* in each group showed no increase in generalization from initial to final test. A possible explanation for this lies in the operation of competing and incompatible response tendencies. Almost all the *Ss* who showed no increase were those who showed no generalization at all during the initial test. It appeared that during that test they formulated for themselves the principle that one should not pull to the test light. Many, in fact, verbalized this principle. Subsequent presentations of the test light during the final test elicited both the principle and the mediating response. Apparently the tendency to react to the principle was stronger than the tendency to react in accordance with the stimulation provided by the mediating response.

Viewed as a whole, however, the study provides at least a tentative answer to the question raised in the introduction and suggests the direction in which future research should go. Figure 1 indicates that the teaching of a common response, in this case the word "Mo," to the training stimulus and stimuli of decreasing similarity from it did raise generalization tendencies at all points. However, it is noteworthy that neither of the two furthest points was raised to the level of the nearest point indicating that there is a relationship between the similarity of two stimuli and the effectiveness of a mediating response. Unfortunately, with the present data it was impossible to make a precise statistical test of this observed interaction between training and test point. A future study using more tests at each point and instructions designed to minimize the buildup of interfering response tendencies should yield data amenable to such analysis. Moreover, since these data do tentatively indicate the presence of the sought after relationship, it would now seem worthwhile to try to discover the precise nature of this relationship. Accordingly, future studies should use stimuli whose similarity to each other can be precisely indicated in j.n.d. units. Also, it would be worthwhile to extend the range of similarity beyond that utilized in this exploratory study in order to determine if a certain minimum amount of similarity is necessary before mediated generalization can be demonstrated.

Summary

Three groups of preschool children were tested for their tendency to generalize a button pushing response from a red light to a reddish orange, an orange and a yellow respectively. One test was given before learning a common name for the training and test stimuli. One test was given afterwards. An increase in generalization following verbal training was observed at all points. The results were discussed in terms of the principle of mediated generalization.

References

1. Birge, Janet. The role of verbal responses in transfer. Unpublished doctor's dissertation, Yale Univer., 1941.
2. Dixon, W. J., & Massey, F. J., Jr. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.
3. Eisman, Bernice S. Attitude formation: The development of a color preference response through mediated generalization. *J. abnorm. soc. Psychol.,* 1955, 50, 321-326.
4. Jeffrey, W. E. The effects of verbal and nonverbal responses in mediating an instrumental act. *J. exp. Psychol.,* 1953, 45, 327-333.
5. Murdock, B. B., Jr. Effects of failure and retroactive inhibition on mediated generalization. *J. exp. Psychol.,* 1952, 44, 156-164.
6. Shipley, Walter C. An apparent transfer of conditioning. *J. gen. Psychol.,* 1933, 8, 382-391.

THE CHILDREN'S FORM OF THE MANIFEST ANXIETY SCALE

Alfred Castaneda,[1] Boyd R. McCandless[1]
Iowa Child Welfare Research Station

and David S. Palermo[1]
Southern Illinois University

In 1951 Taylor (37) published the results of an experiment in which the performance of adults in a classical conditioning situation was found to be related to their scores on a questionnaire relating to anxiety. The questionnaire, now more familiarly known as the Taylor scale of manifest anxiety, was subsequently published (38) with a description of the history of its development and stardardization data. Since then the performance of adults scoring high and low on the scale has been studied in classical conditioning (3, 22, 29, 32, 34, 35, 36, 37), differential conditioning (11, 31, 33), verbal paired associate learning (27), serial anticipation learning (14, 16, 18, 19, 23, 39), motor learning (26, 40), perception (4), maze learning (8, 21), simple and complex reaction time tasks (10, 24, 41) and stimulus generalization (6, 28). In addition, the effects of anxiety have been studied in conjunction with stress (5, 6, 18, 25, 42), failure (19, 26), electric shock (34), incentives (22) and mental set (20). Its relationship to other indices of anxiety, behavior disorders, intelligence, etc., has also been extensively studied (1, 2, 7, 9, 12, 13, 15, 17, 42). A number of these studies have served as the empirical component upon which Spence (30) has based his attempt to clarify the role of motivation in simple and complex learning situations.

In the interest of extending the generality and continued usefulness of this scale over a wider range of different populations, the present study

---

[1] The authors wish to express their thanks and gratitude for the cooperation and assistance received from the following persons: Buford W. Garner, Iowa City Superintendent of Schools; David K. Stewart, Elementary School Curriculum Coordinator; Sylvella Jacobsen, School Psychologist; Donald L. Seavy, Principal, Longfellow School; Margaret Schindhelm, Principal, Henry Sabin School; Albert D. Hensleigh, Principal, Horace Mann School; and Gertrude Smith, Principal, Herbert Hoover School.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

reports the description of and the normative data obtained with a scale of manifest anxiety adapted from Taylor's adult form for use with fourth, fifth and sixth grade children.

An earlier version of the present scale was administered to approximately 60 subjects for the purposes of obtaining information regarding possible difficulties in the instructions for its administration and the comprehensibility of the items. A total of 42 anxiety items were selected and modified, and 11 additional items designed to provide an index of the subject's tendency to falsify his responses to the anxiety items were included in the present form of the test. A similar set of items was included by Taylor from the L scale of the MMPI for the same purpose; hence these 11 items will also be referred to as the L scale. All of these items were then submitted to two elementary school system officials for a final check on their comprehensibility for the population for which they were intended. The 42 anxiety items, constituting the anxiety scale of the present test, are reproduced below with their appropriate ordinal numbers as they appear in the present form of the test. The index of the level of anxiety is obtained by summing the number of these items answered "Yes."

1. It is hard for me to keep my mind on anything.
2. I get nervous when someone watches me work.
3. I feel I have to be best in everything.
4. I blush easily.
6. I notice my heart beats very fast sometimes.
7. At times I feel like shouting.
8. I wish I could be very far from here.
9. Others seem to do things easier than I can.
11. I am secretly afraid of a lot of things.
12. I feel that others do not like the way I do things.
13. I feel alone even when there are people around me.
14. I have trouble making up my mind.
15. I get nervous when things do not go the right way for me.
16. I worry most of the time.
18. I worry about what my parents will say to me.
19. Often I have trouble getting my breath.
20. I get angry easily.
22. My hands feel sweaty.
23. I have to go to the toilet more than most people.
24. Other children are happier than I.
25. I worry about what other people think about me.
26. I have trouble swallowing.
27. I have worried about things that did not really make any difference later.
28. My feelings get hurt easily.
29. I worry about doing the right things.
31. I worry about what is going to happen.
32. It is hard for me to go to sleep at night.
33. I worry about how well I am doing in school.
35. My feelings get hurt easily when I am scolded.
37. I often get lonesome when I am with people.
38. I feel someone will tell me I do things the wrong way.

39. I am afraid of the dark.
40. It is hard for me to keep my mind on my school work.
42. Often I feel sick in my stomach.
43. I worry when I go to bed at night.
44. I often do things I wish I had never done.
45. I get headaches.
46. I often worry about what could happen to my parents.
48. I get tired easily.
50. I have bad dreams.
51. I am nervous.
53. I often worry about something bad happening to me.

The 11 L scale items are reproduced below with their appropriate ordinal numbers as they appear in the present form of the test. Items 10 and 49, if answered "No," contribute to the L scale score as do the remaining nine items if answered "Yes." The index of the subject's tendency to falsify his responses to the anxiety items, then, is the sum of these items answered in the designated manner.

5. I like everyone I know.
10. I would rather win than lose in a game.
17. I am always kind.
21. I always have good manners.
30. I am always good.
34. I am always nice to everyone.
36. I tell the truth every single time.
41. I never get angry.
47. I never say things I shouldn't.
49. It is good to get high grades in school.
52. I never lie.

Copies of this test were distributed to the classroom teacher who administered it to her class on a group basis. The only instructions provided the teacher were those which appeared on the test itself and which she read to the class. The instructions were: "Read each question carefully. Put a circle around the word YES if you think it is true about you. Put a circle around the word NO if you think it is not true about you." Space was provided on the test sheet for the subject to identify himself by name, grade, sex and school. Approximately one week later the classroom teacher readministered the test. A total of 15 classrooms from four different schools participated in the study. A total of 386 children participated in the first administration of the test. However, due primarily to absences, only 361 of these children were tested on the second administration.

ANXIETY SCALE RESULTS

Table 1 presents, by sex as well as by grade, the descriptive statistics for the anxiety scale alone, based on the data obtained on the first administration of the test. A general tendency can be noted for the girls to receive

TABLE 1

ANXIETY SCALE MEANS AND *SD*'S FOR THE DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	Fourth Grade			Fifth Grade			Sixth Grade		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Boys	58	14.97	7.82	71	16.24	7.43	73	16.58	7.39
Girls	62	19.09	7.63	68	17.75	9.23	54	18.50	7.82

higher scores. The results of an analysis of variance performed on these data indicated that the effects of sex were significant at beyond the .005 level, i.e., girls gave significantly more "Yes" responses in comparison with boys. The effects of grade and the interaction between sex and grade were not significant. Plots of the distributions for boys and girls separately revealed them to be highly comparable with the exception of a tendency for the girls' distribution to be somewhat less positively skewed. For the present purposes, however, it seemed sufficient to present only a single distribution based on the scores for both the boys and the girls. This distribution is presented in Figure 1.

Inspection of this distribution reveals a slight positive skew and its general form is comparable to that obtained by Taylor with the adult form of the scale with over 1900 college students. The 50th percentile of the

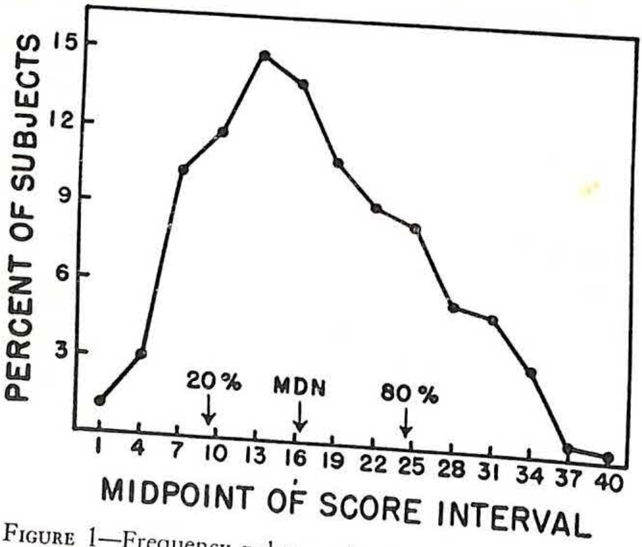


FIGURE 1—Frequency polygon showing percentage of 386 4th, 5th and 6th grade children receiving the indicated scores on the Children's Form of the Manifest Anxiety Scale

TABLE 2

ANXIETY SCALE RETEST CORRELATIONS FOR THE DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	Fourth Grade		Fifth Grade		Sixth Grade	
	N	*r*	*N*	*r*	*N*	*r*
Boys	56	.88	68	.91	65	.82
Girls	61	.70	62	.94	49	.89

distribution falls at about 16, the 80th at about 25 and at about 10 for the 20th percentile. The mean of the distribution is 17.16. When the distributions were plotted for the girls alone the 50th percentile fell at about 18, the 80th at about 27 and at about 11 for the 20th percentile. For the boys these percentiles fell respectively at about 15, 24 and 9. The mean of the distribution for the girls was 18.45 and 15.87 for the boys.

A general picture of the consistency of the anxiety scale as a function of sex as well as grade can be obtained from Table 2. All correlation coefficients (Pearson product-moment) are significant at well beyond the .01 level indicating that the relative position of the subject in his group tended to remain constant over the one-week period.

TABLE 3

L SCALE MEANS AND *SD*'S FOR THE DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	Fourth Grade			Fifth Grade			Sixth Grade		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Boys	43	2.44	1.96	71	1.94	1.97	73	1.84	1.63
Girls	52	2.54	2.01	68	3.07	2.27	54	1.81	1.76

L SCALE RESULTS

Table 3 presents, by grade as well as by sex, the descriptive statistics for the L scale based on the data obtained on the first administration of the test.[2] As with the anxiety scale, a general tendency is again noted for girls to receive higher scores in comparison with boys. Also, there is an apparent tendency for the sixth grade to score lower in comparison with either the fourth or fifth grades. The results of an analysis of variance performed

---

[2] One fifth-grade classroom of 33 was omitted from this and subsequent analyses involving the L scale when it was discovered that additional and misleading instructions had been given regarding the L scale items.

TABLE 4

L SCALE RETEST CORRELATIONS FOR THE DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	Fourth Grade		Fifth Grade		Sixth Grade	
	N	*r*	*N*	*r*	*N*	*r*
Boys	41	.75	68	.69	65	.53
Girls	51	.72	62	.85	49	.61

on these data indicated that the effects of sex were significant at the .025 level and those of grade at the .01 level. The interaction between sex and grade was not significant.

Table 4 provides a general picture of the consistency of the L scale for boys and girls separately as well as for the three different grade levels. All correlations, as with the anxiety scale, are significant at beyond the .01 level. However, in comparison with the anxiety scale, they tend to be somewhat lower on the average. This difference in the magnitude of the correlations undoubtedly reflects the difference in the number of items comprising the two scales.

RELATIONSHIP BETWEEN THE TWO SCALES

Assuming that the tendency, as reflected in the L scale, for any given subject to falsify his responses to the anxiety items could result in a high anxiety score as well as a low one, correlations of a low order between the two scales would be highly desirable.

Table 5 represents the correlations between the two scales, again for the different grade levels and for boys and girls separately. All of these correlation coefficients fail to attain statistical significance and it can be readily noted that with the exception of the correlation for the sixth grade girls they tend to cluster around the zero value.

TABLE 5

CORRELATIONS BETWEEN THE ANXIETY SCALE AND L SCALE FOR THE DIFFERENT GRADE LEVELS AND FOR BOYS AND GIRLS SEPARATELY

	Fourth Grade		Fifth Grade		Sixth Grade	
	N	*r*	*N*	*r*	*N*	*r*
Boys	41	—.04	68	.03	65	—.10
Girls	51	—.01	62	—.11	49	.22

Contributions to the specification of some of the functional properties currently ascribed to anxiety have stemmed from the theoretical and experimental work of a number of psychologists, particularly Mowrer, Hull and Spence. In current behavior theory these properties are those which define drive states. Hence, as it has been explicitly assumed for the adult form of the anxiety scale, the present scale provides a method for measuring the level of drive with the immediate purpose of attempting to determine its role as a determinant of performance in children.

SUMMARY

The present study reports the description of and the normative data obtained with a scale of manifest anxiety adapted from Taylor's adult form for use with fourth, fifth and sixth grade children. A 42-item anxiety scale and an 11-item L scale are described. One-week retest reliabilities averaged at about .90 for the anxiety scale and at about .70 for the L scale. Intercorrelations between the anxiety scale and the L scale clustered around the zero value. Girls were found to score significantly higher than boys on both scales. Significant differences on the L scale were found to be associated with grade.

REFERENCES

1. BECHTOLDT, H. P. Response defined anxiety and MMPI variables. *Proc. Iowa Acad. Sci.*, 1953, 60, 495-499.
2. BENDIG, A. W. Age, sex and the manifest anxiety test. *J. consult. Psychol.*, 1954, 18, 16.
3. BITTERMAN, M. B., & HOLTZMAN, W. H. Conditioning and extinction of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, 47, 615-623.
4. BITTERMAN, M. B., & KNIFFIN, C. W. Manifest anxiety and "perceptual defense." *J. abnorm. soc. Psychol.*, 1953, 48, 248-253.
5. DEESE, J., LAZARUS, R. S., & KEENAN, J. Anxiety, anxiety reduction and stress in learning. *J. exp. Psychol.*, 1953, 46, 55-60.
6. ERIKSEN, C. W. Some personality correlates of stimulus generalization under stress. *J. abnorm. soc. Psychol.*, 1954, 49, 561-565.
7. ERIKSEN, C. W., & DAVIDS, A. The meaning and clinical validity of the Taylor anxiety scale and the hysteria-psychasthenia scales from the MMPI. *J. abnorm. soc. Psychol.*, 1955, 50, 135-137.
8. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, 45, 120-125.

9. Goodstein, L. D. Interrelationships among several measures of anxiety and hostility. *J. consult. Psychol.*, 1954, 18, 35-39.

10. Grice, G. R. Discrimination reaction time as a function of anxiety and intelligence. *J. abnorm. soc. Psychol.*, 1955, 50, 71-74.

11. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, 42, 94-100.

12. Holtzman, W. H., Calvin, A. D., & Bitterman, M. E. New evidence for the validity of Taylor's manifest anxiety scale. *J. abnorm. soc. Psychol.*, 1952, 47, 853-854.

13. Holtzman, W. H., Iscoe, I., & Calvin, A. D. Rorschach color responses and manifest anxiety in college women. *J. consult. Psychol.*, 1954, 18, 317-324.

14. Hughes, J. B., Sprague, J. L., & Bendig, A. W. Anxiety level, response alteration, and performance in serial learning. *J. Psychol.*, 1954, 38, 421-426.

15. Kerrick, Jean S. Some correlates of the Taylor manifest anxiety scale. *J. abnorm. soc. Psychol.*, 1955, 50, 75-77.

16. Laffal, J. Learning and retention of words with associative disturbance. *J. abnorm. soc. Psychol.*, 1952, 47, 454-462.

17. Lauterbach, C. G. Empirical study of the manifest anxiety scale and its relationships to other clinical measures of anxiety. Unpublished doctor's dissertation, State Univer. of Iowa, 1952.

18. Lazarus, R. A., Deese, J., & Hamilton, R. Anxiety and stress in learning: the role of intraserial duplication. *J. exp. Psychol.*, 1954, 47, 111-114.

19. Lucas, J. D. The interactive effects of anxiety, failure, and intraserial duplication. *Amer. J. Psychol.*, 1952, 65, 59-66.

20. Maltzman, I., Fox, J., & Morrisett, L. Some effects of manifest anxiety on mental set. *J. exp. Psychol.*, 1953, 46, 50-54.

21. Matarazzo, J. D., Ulett, G. A., & Saslow, G. Human maze learning performance as a function of increasing levels of anxiety. *J. gen. Psychol.*, in press.

22. Meyer, D. R., Bahrick, H. D., & Fitts, P. M. Incentive, anxiety, and the human blink rate. *J. exp. Psychol.*, 1953, 45, 183-187.

23. Montague, B. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, 45, 91-96.

24. Myers, T. I. Discriminatory reaction performance as a function of test-defined anxiety level and verbal instructions. Unpublished master's thesis, State Univer. of Iowa, 1951.

25. Osler, Sonia F. Intellectual performance as a function of two types of psychological stress. *J. exp. Psychol.*, 1954, 47, 115-121.

26. Price, Helen G. Anxiety and failure as factors in the performance of motor tasks. Unpublished doctor's dissertation, State Univer. of Iowa, 1951.

27. Ramond, C. Anxiety and task as determiners of verbal performance. *J. exp. Psychol.*, 1953, 46, 120-124.

28. Rosenbaum, G. Stimulus generalization as a function of level of experimentally induced anxiety. *J. exp. Psychol.*, 1953, 45, 35-43.

29. Sampson, H., & Bindra, D. "Manifest" anxiety, neurotic anxiety, and the rate of conditioning. *J. abnorm. soc. Psychol.*, 1954, 49, 256-259.

30. Spence, K. W. *Symposium on relationships among learning theory, personality theory and clinical research.* New York: Wiley, 1953.

31. Spence, K. W., & Beecroft, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, 48, 399-403.

32. Spence, K. W., & Farber, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.*, 1953, 45, 116-119.

33. Spence, K. W., & Farber, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, 47, 127-134.

34. Spence, K. W., Farber, I. E., & Taylor, Blaine. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. exp. Psychol.*, 1954, 48, 404-408.

35. Spence, K. W., & Taylor, Janet A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, 42, 183-189.

36. Spence, K. W., & Taylor, Janet A. The relation of conditioned response strength to anxiety in normal, neurotic and psychotic subjects. *J. exp. Psychol.*, 1953, 45, 265-272.

37. Taylor, Janet A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, 41, 81-89.

38. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.

39. Taylor, Janet A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, 44, 61-64.

40. Weatherley, D. A. The effect of anxiety on massed and distributed practice in motor learning. Unpublished master's thesis, Northwestern Univer., 1952.

41. Wenar, C. Reaction time as a function of manifest anxiety and stimulus intensity. *J. abnorm. soc. Psychol.*, 1954, 49, 335-340.

42. Westrope, Martha R. Relations among Rorschach indices, manifest anxiety, and performance under stress. *J. abnorm. soc. Psychol.*, 1953, 48, 129-134.

COMPLEX LEARNING AND PERFORMANCE AS A FUNCTION OF ANXIETY IN CHILDREN AND TASK DIFFICULTY

ALFRED CASTANEDA,
Iowa Child Welfare Research Station

DAVID S. PALERMO
Southern Illinois University

and BOYD R. MCCANDLESS
Iowa Child Welfare Research Station

Within the last few years a number of experiments by Spence (7, 8), Farber (3), Taylor (9, 11) and Montague (5) have attempted to compare the performance of adult anxious and nonanxious *S*s in classical conditioning as well as in more complex forms of learning situations. The presence of anxiety in these *S*s had been defined in terms of their responses on a personality inventory of manifest anxiety developed by Taylor (10). *S*s receiving high scores on this test characteristically exhibit a higher level of performance in the conditioning situation than those receiving low scores. These results have been interpreted to support the assumption that these scores are an index of *S*'s readiness to respond or excitability, which in turn is assumed to reflect their general drive level. Reasoning on the basis of Hull's (4) theoretical formulation relating response strength to drive, i.e., that all habit tendencies aroused by the stimulus situation are multiplied by the total effective drive level then operating, two opposing, though not contradictory, expectations have been derived by Spence and his associates (6) regarding the role of anxiety as a determinant of performance. If, as in classical conditioning, there tends to be but a single or highly dominant response tendency, an increase in the strength of drive should enhance the strength of that response thereby benefiting performance (3, 7, 8, 9). In situations where there is greater likelihood that more than one response tendency is aroused by the stimulus situation, as in complex learning tasks, whether or not an increase in the drive level will aid or impede performance depends on whether or not the criterion (correct) response tendency is stronger (dominant) in relation to other competing tendencies aroused by the stimulus situation. If the correct response tendency is the

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

stronger, performance should be aided by an increase in the drive level. However, should the incorrect response tendency be stronger, the effects of heightened drive should be deleterious to performance. Empirical support for these expectations has also been found in studies dealing with more complex forms of learning in which the Taylor scale has been used as the empirical variable assumed to reflect differences in drive level (3, 5, 11).

The present study is concerned with the performance of fifth grade children on a complex learning task as a function of the relative difficulty of the various components comprising the task and of their scores on a scale of manifest anxiety adapted for children from Taylor's adult form.

METHOD

Subjects

The 37 *S*s in the present study were from among those who participated in the standardization of the children's form of the anxiety scale (1) and who, in addition, participated in a complex learning experiment (2) a year previously. The present study, then, reports the relationship between the anxiety scale scores of these *S*s and their performance in the experiment of the previous year. Although the conventional practice with the adult form is to select *S*s whose scores fall within the upper or lower 20th percentiles of the distribution, the small number of *S*s available in the present study precluded such a procedure. The anxiety scores of the present *S*s ranged from a low of three to a high of 33; hence a score of 18 or above was arbitrarily designated as falling in the high anxious category and a score of 17 or below in the low anxious category. The high anxious group, then, consisted of 21 *S*s, 9 boys and 12 girls, and the low anxious group was composed of 16 *S*s, 6 boys and 10 girls. In addition, the L scale, i.e., that part of the present form of the test assumed to provide an index of *S*'s tendency to falsify his responses to the anxiety items, was not employed. Preliminary analyses based on the L scale indicated that its use would not have appreciably affected the results. In addition, the loss in the number of subjects resulting from designating a criterion of a score of three or less on the L scale would have resulted in a loss of approximately 30 per cent of the subjects from the present sample.

Apparatus

Since a more detailed description of the apparatus may be found in (2), only a brief description of it will be presented here. In essence, it consisted of a rectangular shaped box approximately 9 by 18 by 9 in., painted flat black. A response panel containing five linearly arranged push buttons projected from the box. Centered 3 in. above the response panel was a 1 in. diameter aperture of flashed opal glass. Behind the aperture were five pilot lamps colored either dark red, green, amber, blue or light red. All controls used by *E* were situated at the back of the apparatus. By a simple switching arrangement, *E* could actuate any single light and set any single push button so that depressing it turned off the light. Depression of any other of the four remaining buttons did not affect the light.

Procedure

All *S*s had to learn the same five light button combinations. These five combinations had been previously determined on a random basis. Each single light was presented five times and randomly interspersed within the total 25 presentations. All *S*s were allowed to continue responding until the correct button had been depressed, at which time the next light was presented. *S*s were merely instructed that the task required learning which buttons were associated with which colored lights and in case of an error to select another button until the correct one had been depressed. The particular buttons depressed and the order in which they were depressed for each presentation were recorded for each *S*.

RESULTS AND DISCUSSION

Studies with the adult form of the anxiety scale have indicated that the differential performance of high and low anxious *S*s can differ depending on the particular characteristics of the task. For example, the tendency for the high anxious *S*s to perform more poorly in comparison to low anxious *S*s increases as the difficulty of the task increases (3, 11). Conversely, if the difficulty of the task can be sufficiently decreased, differences in anxiety level may result to the benefit of the high anxious *S*s (5). Hence, these studies indicate that the difficulty existing either among several different tasks (5) or among the various components comprising a given task (3, 11) should be assessed, preferably on some basis independent of the performance of the high and low anxious groups which are being studied. Therefore, in order to determine the possibility that the five light button combinations in the present task may not have been equal with respect to the ease with which they could be learned, 20 *S*s from among those who had participated in the previous experiment, and for whom anxiety scores were not available, were drawn at random and their performance on each of the five combinations was determined. The index of the difficulty of learning a given combination was the number of times, out of five, the first response to the light was the correct one, or more simply, the number of errorless trials.

Table 1 presents the results of this comparison. The combinations are numbered from I to V in order of increasing difficulty, i.e., in order of the decreasing number of errorless trials. For the present purposes it was decided to compare the performance of the high and low anxious *S*s on the two easiest (I and II) and the two most difficult (IV and V) combinations.

TABLE 1

MEAN NUMBER OF ERRORLESS TRIALS MADE BY THE RANDOM SAMPLE OF 20 SUBJECTS ON THE FIVE SEPARATE LIGHT-BUTTON COMBINATIONS

	Combination Number				
	I	II	III	IV	V
M	3.40	3.05	2.45	2.40	2.00
SD	1.48	1.02	1.02	1.62	1.10

Table 2 presents the number of errorless trials for the high and low anxious groups for the easy and difficult combinations separately. It is apparent, on the basis of these data, that the high anxious children performed better, in comparison with the low anxious children, on the easy combinations, but more poorly on the difficult combinations. It can be noted that the performance of the high anxious children appeared to be more affected by the differences in the difficulty of the two sets of combinations than that of the low anxious children.

Table 3 presents a summary of an analysis of variance based on the data presented in Table 2. This indicates that the main effect of anxiety was not significant while the effects of task difficulty were significant at only the .10 level. Of greater interest, however, is the significant interaction between anxiety and task difficulty. This interaction may be interpreted to indicate that the effects of anxiety are dependent on the degree of difficulty involved in the task. This is in accord with the data presented in Table 2, showing that the position of superiority of the high anxious children on the easy combinations, in comparison with the low anxious children, is completely reversed on the more difficult combinations. Tests of the simple effects indicated that only the difference between the low and high anxious children on the difficult combinations was significant at beyond the .05 level ($F = 4.99$, $df = 1, 35$).

TABLE 2

MEAN NUMBER OF ERRORLESS TRIALS ON THE EASY AND DIFFICULT COMBINATIONS FOR THE HIGH AND LOW ANXIOUS GROUPS SEPARATELY

	Difficulty Level			
	Easy		Hard	
Group	*M*	*SD*	*M*	*SD*
High Anxious	5.00	2.41	3.33	1.29
Low Anxious	4.38	2.13	4.81	2.66

TABLE 3

SUMMARY OF ANALYSIS OF VARIANCE BASED ON NUMBER OF ERRORLESS TRIALS

Source	*df*	*MS*	*F*	*p*
Anxiety	1	3.31	....	
Error (b)	35	6.23		
Difficulty	1	10.59	2.99	$<.10$
Anxiety × Difficulty	1	20.10	5.68	$<.025$
Error (w)	35	3.54		

The present results are in general agreement with those found with the adult form of the anxiety scale in similar learning situations. In view of the fact that a difference in a scale score of only one was used as the basis for differentiating the low from the high anxious children in lieu of the more conventional procedure of selecting *S*s from the extreme ends of the distribution, and that there was approximately a one-year interval between participation in the experiment and the time the anxiety scores were obtained, these results offer some encouragement to any future use of the scale with children.

Assuming that the index of the difficulty of the two sets of combinations reflected the presence of competing response tendencies of increasing strength, the present results indicated that the effects of anxiety varied as a function of this index. Consistent with the theoretical formulations outlined earlier, it could be assumed that on the easy combinations the correct response was stronger in relation to other competing responses which may have been aroused by the stimulus elements of these combinations. Under such conditions the effects of high anxiety should be expected to be beneficial to performance. In the case of the difficult combinations it could be assumed that other competing response tendencies were stronger in relation to the correct one; hence the effects of high anxiety, according to the theory, should be expected to be deleterious to performance.

In general these results suggest that the presence of anxiety in children can serve to determine performance in much the same way as with adults. Of more general importance, perhaps, the present results support the notion that the effects of anxiety can be more profitably studied if the characteristics of the task can be specified with regard to the number and relative strength of the competing response tendencies that may be aroused, as well as to whether or not the correct response is dominant. This appears to be in accord with clinical and casual observation that the consequences of a heightened level of anxiety may vary with the particular stimulating circumstances and the particular responses the individual has learned to make in such situations on the basis of previous experience.

SUMMARY

The present study was concerned with the performance of fifth grade children in a complex learning task as a function of their scores on a scale of manifest anxiety adapted from Taylor's adult form for use with children and the relative difficulty of the various components comprising the task.

The high anxious group was composed of 21 children and the low anxious group of 16. The relative difficulty existing among the five components comprising the task was determined on the basis of performance of an independent sample of *S*s. A statistically significant interaction was found between anxiety and task difficulty. This interaction was based on the tendency for the performance of the high anxious children to be inferior, in comparison with the low anxious children, on the difficult components of the task but with a tendency for their performance to be superior on the less difficult components.

REFERENCES

1. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Develpm.,* 1956, 27, 317-326.
2. CASTANEDA, A., & PALERMO, D. S. Psychomotor performance as a function of amount of training and stress. *J. exp. Psychol.,* 1955, 50, 175-179.
3. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.,* 1953, 45, 120-125.
4. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
5. MONTAGUE, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.,* 1953, 45, 91-96.
6. SPENCE, K. W. *Symposium on relationships among learning theory, personality theory, and clinical research.* New York: Wiley, 1953.
7. SPENCE, K. W., & FARBER, I. E. Conditioning and extinction as a function of anxiety. *J. exp. Psychol.,* 1953, 45, 116-119.
8. SPENCE, K. W., & TAYLOR, J. A. Anxiety and the strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.,* 1951, 42, 183-188.
9. TAYLOR, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.,* 1951, 41, 81-92.
10. TAYLOR, J. A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.,* 1953, 48, 285-290.
11. TAYLOR, J. A., & SPENCE, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.,* 1952, 44, 61-64.

THE RELATIONSHIP OF ANXIETY IN CHILDREN TO PERFORMANCE IN A COMPLEX LEARNING TASK

David S. Palermo,
Southern Illinois University

Alfred Castaneda and Boyd R. McCandless
Iowa Child Welfare Research Station

Using the Taylor scale of manifest anxiety (8) as a measure of motivation, a number of studies (4, 6, 7, 9) have attempted to relate the effects of different motivational levels to the performance of adult *S*s in a variety of learning situations. The results of these experiments suggest that the performance of high anxious *S*s, when compared with low anxious *S*s, is superior in simple learning situations such as classical conditioning but inferior in more complex tasks such as trial-and-error learning.

In an attempt to broaden the use of the Taylor scale, Castaneda, McCandless, and Palermo (1) have recently revised the scale so that it is applicable to children in the fourth, fifth, and sixth grades. It was felt that problems similar in nature to those investigated with adults might be studied with children if such a scale were available.

According to the theory based on Hull (3) and extended by Spence (5), the effects of increases in motivational level on performance depend upon the relative strength of the correct and the incorrect or competing responses. If the correct response is dominant (strongest) over those incorrect responses competing with it or if it is the only response aroused in the situation, then an increase in motivation should facilitate performance. Thus, in the case of eyelid conditioning, if it is assumed that the correct response is dominant and there are few competing responses, it would be expected that an increase in motivation would lead to superior performance (7).

However, in more complex learning situations where one or more competing incorrect responses are dominant, the theory would predict an impairment in performance as a result of an increase in motivation. The

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

present investigation is concerned with attempting to determine the relationship of anxiety or motivation in children to performance in such a situation. *S*s were required to learn to turn off different colored lights by the buttons which were connected to them. In this situation, if the correct response is not dominant, it would be expected that an increase in motivation should lead to poorer performance.

METHOD

Subjects

A total of 36 *S*s from the standardization population of the children's form of the Taylor scale (8) were used as *S*s. These *S*s were selected from four fourth grade classes in three Iowa City public schools[1] on the basis of extreme scores on the scale. The anxious group consisted of 9 girls and 9 boys drawn at random from those *S*s who participated in this study and whose scores fell approximately in the upper 20 per cent of the total standardization group, while the nonanxious group was composed of 9 girls and 9 boys drawn at random from those *S*s who participated in this study and falling approximately in the lower 20 per cent. Anxiety scores for the nonanxious group ranged form 3 to 11 and from 23 to 33 for the anxious *S*s.

Apparatus and Procedure

The apparatus consisted of a box 13 by 12½ by 5½ in. painted flat black. Projecting from its base was a sloping response panel approximately 7½ by 9½ by 13 in. Two push buttons were arranged on this panel on a horizontal plane spaced 6 in. apart and 2 in. from *S*'s edge of the panel and toward its center. On the panel between, but 2½ in. above, the two push buttons was a green jeweled reflector illuminated by a 6.3 volt pilot lamp. Centered 4 in. above the panel was a 1 in. diameter aperture of flashed opal glass. Housed behind the aperture were four colored lights: red, blue, amber, and white. The construction of the apparatus was such that by means of a rotary selector switch any single light could be activated and by means of a toggle switch either one of the two buttons could be set to turn it off. If the correct button was selected, the light went off and the green light on the response panel was activated. Incorrect responses affected neither the stimulus light nor the green light on the response panel.

All *S*s were given 20 trials. All of the lights did not appear with equal frequency but the order was random except that no single light appeared twice in succession. Each light remained on until the correct response was made. Two of the lights were turned out by one button and two were turned out by the other button.

---

[1] The authors wish to express their appreciation to David K. Stewart, Elementary School Curriculum Coordinator in the Iowa City School System, and the principals and teachers involved for their cooperation and assistance.

The instructions were designed to indicate to *S* that he was to learn which button turned off each light and that if an error was made correction would be allowed.

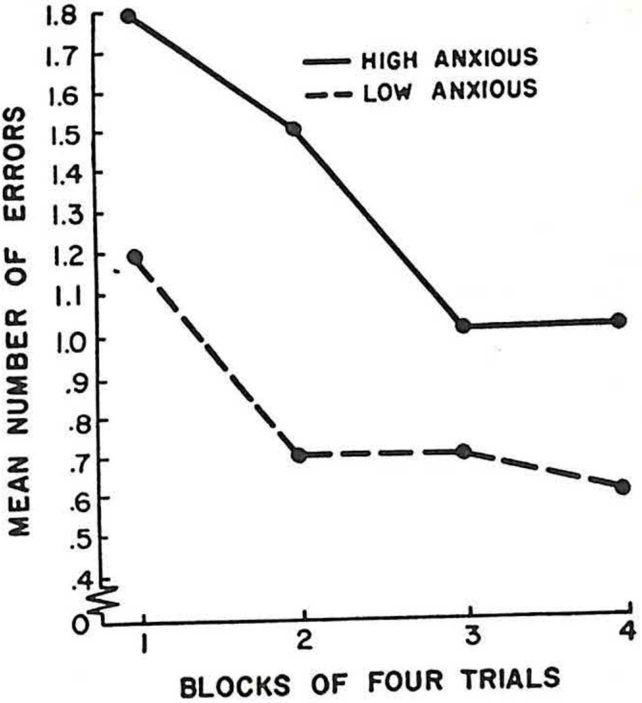


Figure 1—Error curves for the anxious and nonanxious *Ss* plotted in blocks of five trials.

Results and Discussion

Figure 1 presents the learning curves of the anxious and the nonanxious *Ss* in terms of errors in blocks of five trials. It will be observed that with respect to each block of trials the nonanxious *Ss* were superior to the anxious *Ss*.

A summary of an analysis of variance based upon these data is presented in Table 1. As may be seen there is a significant decrease in errors over blocks of trials. More important, however, is the fact that the over-all differences between the errors made by the anxious and nonanxious groups is significant.

Previous experiments using the adult form of the Taylor scale have indicated that in complex learning situations where the dominant response

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE FOR ERRORS IN BLOCKS OF FIVE TRIALS

Source	*df*	*MS*	*F*	*p*
Anxiety	1	12.25	4.71	$< .05$
Error (b)	34	2.60		
Trials	3	3.13	5.49	$< .005$
Trials × Anxiety	3	.42	....	
Error (w)	102	.57		

is not correct, the anxious (high motivation) *S*s show inferior performance to that of the nonanxious (low motivation) *S*s. The present results are in line with previous work suggesting that the theory applied to adult *S*s is applicable in the present study with children and that the children's anxiety scale is a useful measure of motivational level.

As was stated previously, according to the theory the effects of increases in motivation on performance depend upon the relative strength of the correct and incorrect responses aroused by the experimental situation. The results of the present study and those of previous studies (2, 4, 6, 7, 9) suggest that specific attention should be given to the performance of anxious and nonanxious *S*s in situations where the strength of the correct and incorrect responses may be determined. For example, with problems such as the present one, it might be expected that the differences between the anxious and nonanxious *S*s would increase with increases in the number of light-button combinations since there would be a greater probability that more strong competing responses would be aroused. In discrimination learning problems it would be expected that difficult discriminations would be easier for the nonanxious than for the anxious. In the area of verbal learning it might be expected that a list of words of high association value and low intralist similarity would be easier for anxious *S*s than for nonanxious *S*s (4).

SUMMARY

The present experiment was concerned with the relation of motivational level to performance in a trial-and-error learning situation. The specific theoretical implication investigated was that the performance of anxious (high motivation) *S*s would be inferior to that of nonanxious (low motivation) *S*s in a learning situation which involved the presence of competing incorrect responses.

A group of 36 fourth grade *S*s were chosen on the basis of extreme scores made on a test of manifest anxiety, the 18 *S*s with high scores being

the anxious group and the 18 *Ss* with low scores being the nonanxious group. Both groups were run in a light-button motor task involving the presentation of 20 choices between two buttons.

The results were in agreement with similar studies using adult *Ss* and with the theoretical expectation that the anxious (high motivation) *Ss* would make significantly more errors in the learning task.

REFERENCES

1. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, 27, 317-326.
2. CASTANEDA, A., PALERMO, D. S., & McCANDLESS, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Develpm.*, 1956, 27, 327-332.
3. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
4. MONTAGUE, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, 45, 91-96.
5. SPENCE, K. W. *Symposium on relationships among learning theory, personality theory and clinical research.* New York: Wiley, 1953.
6. SPENCE, K. W., & FARBER, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, 47, 127-134.
7. TAYLOR, JANET A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, 41, 81-89.
8. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.
9. TAYLOR, JANET A., & SPENCE, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, 44, 61-64.

VERBAL SOLUTIONS TO PARENT-CHILD PROBLEMS[1]

PHILIP W. JACKSON
The University of Chicago

Children frequently cross the adult-defined boundaries which separate desirable and undesirable behavior. When this occurs, adults, especially parents, often take action aimed at the reduction of these "boundary-crossings." Study of this action promises to yield valuable information concerning the parent's concept of the parental role, his basic attitudes toward the use of authority, the nature of the child, etc. The importance of this action from the standpoint of its effect upon children is obvious. It is hardly surprising then that some of the earliest studies in the area of child development concerned an analysis of parental control and of children's attitudes toward that control (12).

Despite the importance of the problem, empirical investigations of parental reaction to a child's misbehavior have been less than successful. The obvious difficulty which confronts the researcher is the unpremeditated or spontaneous character of these events. He just cannot be there when it happens. If he is fortunate enough to be present, his very presence represents a variable which may affect quite drastically the events he hopes to observe. As a substitute for this "on the spot" observing, researchers have most frequently used two approaches. They have turned either to reports of parents or children concerning past parental action (4, 9, 14) or to some assessment of parental attitudes toward children (2, 3, 6, 10, 11, 13). The techniques used to accumulate this information have ranged from survey questionnaires to clinical interviews. Conclusions based upon data provided

---

[1] This research is part of a study done at Teachers College, Columbia University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to acknowledge his deep indebtedness to Professors Millie Almy, Arthur I. Gates, Arthur T. Jersild, and Irving Lorge of that institution for their many valuable suggestions and criticisms.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

by these approaches assume that some relationship exists between the parent's verbal report and his action in the home.

That the two approaches cited above have been fruitful is clear. That they are not without serious limitations should be equally clear. It is imperative then that alternative approaches be used to verify the conclusions culled from parental reports and attitude studies. One such alternative might be to present the parent with a hypothetical problem situation which he is asked to "solve." This little-used technique provided the body of data reported in this paper.

It is the purpose of this paper (*a*) to describe a system of coding by which the free responses of a group of adults to a series of hypothetical parent-child problems were classified, and (*b*) to report the substantive findings related to sex differences. Conclusions based upon these findings are compared with similar conclusions arrived at through different approaches.

Subjects and Procedures

The subjects used in this study were 105 parents of freshmen and sophomore college students. Forty-five fathers and 60 mothers participated in the study. The mean age of the fathers was 45.2; of the mothers, 43.1. Socioeconomically, the group was judged to consist largely of middle and lower class parents. This judgement was based on information concerning the occupations and education of the subjects.

A questionnaire-type instrument, designed by the writer, was used to obtain the data of this study. The instrument was sent to the parents by mail. An accompanying letter requested the subjects to omit any identifying information in their responses. Although the parents believed their responses to be anonymous, identifying marks hidden in the body of the instrument provided the researcher with the identity of the respondent. Briefly, the subjects were asked to write free responses to the following situations:

> Situation 1: You are seated in your living room when your twelve-year-old son enters. As he takes off his jacket, a pack of cigarettes falls from his pocket.
>
> Situation 2: You return from a shopping trip one afternoon, and discover that your six-year-old daughter has cut off her beautiful curls.
>
> Situation 3: You ask your adolescent son what time he came home last night. He tells you midnight, but you had been downstairs until one o'clock in the morning, so you know this is not true.
>
> Situation 4: Looking out of the window, you see your seven-year-old daughter with some other children tying a string of cans to a dog's tail.
>
> Situation 5: While playing with his building set, your five-year-old son hits his thumb with a hammer and he begins cursing loudly with fairly vile language.
>
> Situation 6: Glancing into your bedroom, you notice your twelve-year-old daughter taking money from your purse.

Situation 7: After a delicious dinner is placed before him, your six-year-old son says he is not hungry.

Situation 8: You tell your four-year-old that he cannot have any candy before supper. He says, "I want some," and begins to strike you with his fists.

Situation 9: After telling your ten-year-old son that he can't go swimming with the boys, you hear him mumble a nasty description about you.

Situation 10: When told that one of her teachers has been seriously injured in a car accident, your thirteen-year-old daughter remarks, "That's good. I didn't like her anyway."

Situation 11: You have asked your sixteen-year-old daughter not to associate with a certain boy in her class. She tells you she is going to spend the evening at a girl friend's house, but, while running an errand in town, you see her stepping into a car with that boy.

Although no attempt was made to determine the real-life frequency of the situations depicted, it was assumed that each situation would appear somewhat realistic to the subject. This assumption was confirmed by the many comments of the this-is-exactly-what-happened-to-me variety which a group of college students made during a trial run of the instrument.

Codification of the Responses[2]

Before the responses were examined, a trial set of categories was constructed. These categories resulted from an examination of the responses accumulated in the trial run mentioned above. The final categories were, with little change, those culled from the trial data.

An outline of the major categories used in the final classification appears below. For a more detailed listing of the categories the reader is referred to (7).

I. *Handling of Situation*

A. Acceptance
  1. Accept because nothing is wrong with behavior.
  2. Accept because a child did it.
  3. Accept because of infrequency.
  4. Accept under limited conditions.
  5. Accept to prevent deceit.

B. Explanation and/or Discussion
  1. Evaluate behavior and give reasons.
  2. Define conditions under which behavior might occur.
  3. Explain broad principles to child.
  4. Ask "loaded" question.

C. Natural Consequences
  1. Child will suffer enough from his own activities.

D. Information Giving
  1. Let child know that parent was aware of what was done.
  2. Evaluate behavior without explanation.

---

[2] The author is particularly indebted to Professor Irving Lorge for the many valuable suggestions which he offered at this stage of the study.

Bureau Edl. Psy. Research
TRAINING COLLEGE
Dated
Acc. No

E. Bribery
   1. Promise reward for deterring.
   2. Call child's attention to a future pleasure.
F. Reprimand
   1. Scold, berate, argue.
   2. "Blow my top," "holler."
G. Assertion of Authority
   1. Urge the child "not to."
   2. Divert attention.
   3. Order, forbid, demand.
   4. Force to stop.
   5. Supervise child more closely.
   6. Alter situation so that it won't happen again.
H. Threat
   1. Of deprivation.
   2. Of natural consequences.
   3. Of physical punishment.
   4. Of loss of status, love.
   5. Of punishment (not elsewhere considered).
I. Compensation
   1. Have child do extra work.
   2. Have child pray for forgiveness.
J. Deprivation
   1. Remove privileges.
   2. Limit action, e.g., "Call indoors."
   3. Deprive child of social contact by ignoring.
K. Coercion to Repeat Behavior
   1. Noncorrection to get repeated natural consequences.
   2. Noncorrection to lead to guilt, shame.
L. Punishment (not otherwise specified)
M. Forced Admission of Error or Guilt
   1. Get word of honor.
   2. Get child to apologize.
N. Physical Punishment
   1. Spank, slap, whip, etc.
   2. Other, e.g. "Wash out mouth with soap."
O. Retribution
   1. Pay back in kind.
P. Creation of Shame
   1. Ridicule.
   2. Psychological pain, e.g. "Make child feel bad."
Q. Creation of Fear
   1. Create fear of natural consequences.
   2. Create fear of punishment.

II. *Elaboration of Situation*

X. Addition of Information
   1. Supplies facts (probable or definite).
   2. Supplies causes (probable or definite).
   3. Makes general statement, e.g. "All children do this."
Y. Search for Information
   1. Seeks facts.

2. Seeks causes.
3. Seeks general information, e.g. "I wonder if all children go through this state."

Z. Avoidance of Situation

1. Denies possibility of situation.
2. Expresses disapproval of behavior but does not act.
3. Places responsibility for action upon some other person.
4. Gives ambiguous reply, e.g. "So what?"

III. *Negation of Situation*

(In cases where the subject reported that he would *not* handle the situation in a given manner, the letter representing the method was placed in brackets with the following superscripts used to denote the reason for rejecting the method.)

( ) ^{0} No reason given.
( ) ^{1} Behavior not considered serious and/or frequent enough.
( ) ^{2} Method would be ineffective.
( ) ^{3} Method not appropriate to age of child.
( ) ^{4} Other.

Using the above categories it was possible to classify the major elements of extremely elaborate responses without wasting much of their richness. Due to space limitations, only two examples of this coding procedure will be given.

Response to situation 9: "First of all, this wouldn't happen to my son because I'd teach him to have a civil tongue. But if it did happen I would make him say the description out loud—he would feel pretty foolish. Then I would get the boy to come into the house and he would get the Ten Commandments read to him, especially 'Honor thy father and thy mother.'" Coded response: $Z_1K_2J_2B_3$.

Response to Situation 8: "I'd yell at him for hitting me but I probably wouldn't spank him since that might make matters worse. Then I'd explain to him that eating candy before dinner would spoil his appetite. I'd also promise him that he could have some candy after dinner if he stopped yelling." Coded response: $F_2(N_1)^2B_1E_2$.

In addition to the writer, an independent judge coded the responses of 20 subjects to the first four parent-child situations. The independent judge found 203 elements in these 80 responses as compared with 214 elements found by the writer. The two judgments were in perfect agreement on 181 or 84 per cent of the response elements.

Quantification of Coded Responses

As judged by five independent child psychologists, the methods of handling the situations are presented above in rank order: beginning with those which might place little pressure toward conformity upon the child and ending with those methods which might place great pressure toward conformity upon the child. The five psychologists ranked the methods on the basis of the amount of pressure toward conformity which they abstractly associated with each method. The coefficient of concordance (8) among the five independent judgments was .93. Using this agreement as a rationale, the writer conceptualized a continuum of coercion upon which these methods were thought to be located. Thus, Bribery would be considered "intrinsically" more coercive than Information Giving; Reprimanding would be more coercive than Bribery, etc. Although it was impossible to locate logical steps on this continuum of coercion, once the continuum was defined some quantitative approximation of "degree" or "amount" seemed justified. Lacking the conditions which would make for a more sophisticated quantification, consecutive numbers were assigned to the ranked methods. Following this procedure Method A (Acceptance) was given the value of 1, Method B (Explanation) a value of 2, etc. These values were identical with the rank values given the methods by the group of psychologists.

Once the methods were assigned numerical values the possibility of comparing the "coerciveness" of two responses seemed apparent. This possibility was complicated, however, by the fact that most subjects suggested the use of more than one method of control in their response to a single situation. Thus, Subject X in response to Situation 1 might suggest Information Giving, Deprivation, and Threat. The problem then was to find some way of comparing X's response with that of Subject Y who in response to the same situation suggested the use of Physical Punishment and Explanation. As a solution the response of each subject was represented by two "scores." The numerical value of the *most* coercive method suggested in the response was called "high coercion score." The corresponding value of the *least* coercive method in the response was called "low coercion score." By comparing the high coercion score of two subjects it was possible to determine which of them suggested the most coercive method in response to a particular situation. A comparison of their low coercion scores indicated which of the two subjects suggested the milder method of control. By summing the high or low coercion scores of a number of subjects it was possible to compare groups in the same fashion as described above.

A third value, the difference between the high and low coercion scores of a subject, was given the term "coercion range." A subject who suggested both very severe and very mild methods obviously would have a larger coercion range than would a subject whose suggested methods clustered around the middle of the continuum of coercion.

SUBSTANTIVE FINDINGS

The findings of this study which deal with differences between fathers and mothers are twofold.

TABLE I

MEAN HIGH AND MEAN LOW COERCION SCORES OF FATHERS AND MOTHERS

	Mean High Score		*Mean Low Score*	
Parent-Child Situation	Fathers ($N=55$)	Mothers ($N=60$)	Fathers ($N=55$)	Mothers ($N=60$)
1	6.30	7.95	2.87	2.58
2	6.57	7.70	3.17	2.80
3	7.53	8.74	5.31	4.40
4	6.46	8.19	3.54	2.76
5	8.26	8.57	5.49	4.11
6	6.75	7.13	2.96	3.76
7	7.03	6.85	4.45	5.09
8	10.77	11.48	8.69	6.04
9	8.83	8.51	5.77	4.64
10	3.61	4.28	2.70	2.60
11	8.32	8.01	3.94	3.54

NOTE.—Considering all situations, the mean high coercion score of mothers was significantly greater than that of fathers ($p<.05$). Differences in mean low coercion scores were not statistically significant.

First, mothers suggested methods of control that were more coercive than were those suggested by fathers. Comparison of the high coercion scores of these two groups showed the mothers to be more coercive than were the fathers in response to eight of the eleven situations. These differences are observable in Table I by comparing parallel cells in the first two columns. Using an analysis of variance these differences yielded an F ratio of 6.41 ($F_{95} = 4.96$).

While there was a tendency for the low coercion scores of women to be lower than those for men, this difference was not statistically significant. The means of the low coercion scores for mothers and fathers are given in Table I.

Secondly, in their responses to parent-child situations, mothers vacillated more between mild and severe methods than did fathers. A comparison of the coercion ranges of these two groups showed this value to be greater for mothers than for fathers in the responses to nine of the eleven situations. The size and direction of these differences are observable in Table 2. Using an analysis of variance these differences yielded an F ratio of 10.17 ($F_{99} = 10.04$).

In sum, the above findings suggest that mothers subscribe to the use of methods of control which are more coercive than those suggested by fathers. However, mothers are also more likely than are fathers to couple these highly coercive methods with some of the milder methods of control.

TABLE 2

COERCION RANGES OF FATHERS AND MOTHERS

	Coercion Range of	
Parent-Child Situation	Fathers ($N=55$)	Mothers ($N=60$)
1	3.43	5.37
2	3.40	4.90
3	2.22	4.34
4	2.92	5.43
5	2.77	4.46
6	3.79	3.37
7	2.58	1.76
8	2.08	5.44
9	3.06	3.87
10	.91	1.68
11	4.38	4.47

NOTE.—Considering all situations, the coercion range was larger for mothers than for fathers ($p<.01$).

DISCUSSION

Few studies which concern parental action or parental attitudes fail to mention differences between the sexes. In general, the findings portray fathers as being more punitive, restrictive or authoritarian than are mothers, who tend toward the "warm," "permissive" end of the continuum. Although recent studies suggest that the fathers of today are less punitive than were their prototypes of a generation ago, they also suggest that both parents are moving toward the permissive end of the continuum without drastically changing their positions relative to each other (9). These findings agree in substance with the widely-held cultural standards of femininity and masculinity. Thus, the notion of a punitive, aggressive male and a loving, passive female are at least not contradicted by research findings. The results of this study, however, raise some questions concerning these traditional sex differences. Obviously, the findings reported above cannot be directly compared with the results of previous studies due to differences in the nature of the data. The conclusions, however, which seem to follow from these data are at variance with conclusions which were derived from other studies dealing with the general problem of parental behavior.

Although it would be interesting to speculate on why females are seen as being less punitive than are males when the measuring instrument is some type of an attitude questionnaire, no attempt will be made to do so here. An attempt will be made, however, to explain why this difference seems to reverse itself when the object of investigation is the written response of a parent to a hypothetical parent-child situation.

To offer an *a posteriori* explanation for the reported sex differences, it is necessary to consider the expectations which our society places upon the behavior of each sex in general as well as the differential role in child training which is expected of men and women. Considering the latter expectation first, it would seem that the mother is held more responsible for the child's behavior than is the father. This expectation results, in part, from the father's prolonged absence from the home as well as from his involvement in the more general task of "caring for" the family. Because of their reflection on her performance as a mother, acts of misbehavior would be more important to her than they would be to the father. The mother then would be more apt to use methods of control which would ensure the disappearance of the undesired behavior. Hence, we find in her solutions to hypothetical parent-child problems the suggested use of methods of control which are more coercive than those suggested by males.

However, in our society punitive or aggressive behavior is tolerated more in men than in women. The cultural concept of a punitive father and a warm, loving mother is one that tends to persist. Therefore, a woman is violating the societal concept of femininity to the degree that she behaves in an aggressive or punitive manner. Within the context of the parent-child relationship the pressing demands of being a "good mother" frequently may necessitate the use of methods which are decidedly aggressive. The mother may dilute these unfeminine actions, however, by coupling them with milder, more feminine methods. Thus it was found that in their solutions to parent-child problems women coupled severe methods with mild methods more so than did men.

The above explanation is in accord with recent formulations in role conflict theory. A concise statement of this theory follows:

> In certain situations *role conflicts* occur. That is, the situations are so ordered that an actor is required to fill simultaneously two or more roles that present inconsistent, contradictory, or even mutually exclusive expectations. The actor cannot realistically conform to these expectations. He is then forced to choose one of several alternatives: he may abandon one role and cling to the other, he may attempt some compromise between the roles, or he may withdraw either physically or psychologically from the roles altogether (5, p. 165).

Stated in terms of this study, the female respondents found it difficult to fill the roles of "mother" and "woman" simultaneously. Of the alternatives suggested in the above quotation, they predominately chose the second. Their attempt at a compromise between the demands of femininity and the demands of motherhood resulted in the patterning of responses described previously.

A final word should be said about the multiplicity of methods suggested in response to single situations. If these responses are at all related to the real life action of parents, this finding is important. It suggests that a one-to-one relationship between a problem situation and a method of control rarely exists. Parents are quite versatile. They do not threaten *or* scold *or* spank; rather, they are more likely to threaten *and* scold *and* spank. One possible reason for this multiplicity of action was discussed above. If it occurs in real life situations, the effect upon the child of this buck-shot approach remains a question worthy of further research.

SUMMARY

A content analysis technique was used to classify the written responses of 167 college students and 105 parents of these students to 11 hypothetical parent-child situations. A procedure for quantifying the classified responses was described.

Differences in the responses of fathers and mothers were found to be at variance with the popular stereotypes of a punitive male and a permissive female. Tentative explanations of these differences were offered within the framework of role conflict theory.

REFERENCES

1. BEHRENS, MARJORIE L. Child rearing and the character structure of the mother. *Child Develpm.,* 1954, 25, 225-238.
2. BLOCK, J. Personality characteristics associated with fathers' attitudes toward child rearing. *Child Develpm.,* 1955, 26, 41-48.
3. COAST, LOUISE C. A study of the knowledge and attitudes of parents and pre-school children. Researches in Parent Education, IV. *Univ. Iowa Stud. Child Welfare,* 1939, 17, 159-181.
4. CUTTS, NORMA E., & MOSELEY N. *Better home discipline.* New York: Appleton-Century-Crofts, 1952.
5. GETZELS, J. W., & GUBA, E. G. Role, role conflict, and effectiveness: an empirical study. *Amer. sociol. Rev.,* 1954, 19, 164-175.
6. JACK, LOIS M. A device for the measurement of parent attitudes and practices. *Univ. Iowa Stud. Child Welfare,* 1932, 6, 135-149.
7. JACKSON, P. W. Verbal solutions to parent-child problems and reports of experiences with punishment. Unpublished doctor's dissertation, Columbia Univer., 1955.
8. KENDALL, M. G. *Rank correlation methods.* London: C. Griffin, 1948.
9. RADKE, MARIAN J. *The relattion of parental authority to children's behavior and attitudes.* Minneapolis: Univ. of Minnesota Press, 1946.

10. SHOBEN, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.

11. STAPLES, RUTH, & SMITH, JUNE WARDEN. Attitudes of grandmothers and mothers toward child-rearing practices. *Child Develpm.*, 1954, 25, 91-97.

12. STOGDILL, R. M. Experiments in the measurement of attitudes toward children: 1899-1935. *Child Develpm.*, 1936, 1, 31-36.

13. STOGDILL, R. M. The measurement of attitudes toward parental control. *J. appl. Psychol.*, 1936, 20, 359-367.

14. STOTT, L. H. Home punishment of adolescent. *J. genet. Psychol.*, 1940, 57, 415-428.

REFERENCE GROUPS OF RURAL HIGH SCHOOL YOUTH[1]

ARNOLD M. ROSE
University of Minnesota

The hypotheses of this paper are the ones long held in sociology that, in a pluralistic society such as our own, an individual is moved to rank the groups to which he belongs or which he otherwise knows about and that his attitudes and behavior reflect the dominant values of the top groups in this hierarchy more than those of the lower groups. Recently this concept of the highly-valued groups has been given the label "reference groups," a term apparently used first by Herbert Hyman (1). This paper considers only those reference groups which are also membership groups (that is, groups to which the individuals studied themselves belong).

The data consists of answers to questionnaires filled out by all students (except for those absent on the typical school day when the survey was taken) in four rural high schools, representing widely different areas in Minnesota. The questionnaires were administered in classrooms by an advanced graduate student at the University of Minnesota, who assured the subjects of anonymity and of the legitimate purposes of the study.

The main question used to ascertain reference groups was "In your life which is most important?" and permitted checking of the following: (*a*) school chums; (*b*) relatives (uncles, aunts and cousins); (*c*) social clubs; (*d*) work groups; (*e*) church groups; (*f*) immediate family (father, mother, brothers, sisters). Of the 582 students filling out the questionnaire, 18 provided no answer to this question and 54 gave more than one answer; both of these categories of individuals were excluded from the analysis. Only four students gave "relatives" as their reference group; the small

---

[1] Design and analysis of this study were carried on under a grant from the Rockefeller Foundation, and the field work was financed by the Minnesota Agricultural Experiment Station and the Graduate School of the University of Minnesota. Field work was done by Edward Hassinger and William Cofell as part of a larger study under the direction of Douglas G. Marshall and Lowry Nelson. The author expresses his appreciation to all the above as well as to Marvin J. Taves and Caroline B. Rose for a critical reading of this paper.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

TABLE I

BACKGROUND OF CHOICE OF REFERENCE GROUPS

	PERCENTAGE GIVING INDICATED ANSWER AMONG:					
	Boys for whom following groups are most important			Girls for whom following groups are most important		
	School Chums	*Immediate Family*	*Organized Groups*	*School Chums*	*Immediate Family*	*Organized Groups*
Reported life satisfaction of father						
Very satisfied	29.4	41.0	39.6	46.7	42.6	31.6
Satisfied	53.0	38.5	43.7	30.0	30.2	47.3
Average	17.6	13.9	12.5	13.3	12.9	15.8
Dissatisfied	0.0	0.6	0.0	0.0	2.7	0.0
Very dissatisfied ..	0.0	0.0	2.1	3.3	0.0	0.0
No answer	0.0	6.0	2.1	6.7	11.6	5.3
Reported happiness of mother						
Very happy	41.2	34.9	33.3	35.7	33.8	36.8
Happy	29.4	49.5	41.7	44.3	46.2	42.1
Average	29.4	10.2	14.6	13.3	16.0	10.5
Unhappy	0.0	3.0	2.1	6.7	1.3	5.3
Very unhappy	0.0	0.6	0.0	0.0	0.0	0.0
No answer	0.0	1.8	8.3	0.0	2.7	5.3
Father's participation in voluntary associations						
Yes	58.8	48.2	37.5	30.0	43.1	42.1
No	41.2	45.8	58.3	63.3	49.3	52.5
No answer	0.0	6.0	4.2	6.7	7.6	5.4
Mother's participation in voluntary associations						
Yes	58.8	62.0	47.9	43.3	40.9	47.4
No	41.2	35.5	47.9	56.7	57.3	47.4
No answer	0.0	2.5	4.2	0.0	1.8	5.2
Number of siblings						
None or one	17.6	10.3	10.4	20.6	12.5	16.0
Two or three	17.6	37.9	31.3	41.0	29.5	5.3
Four to seven	53.0	37.9	35.4	31.6	42.6	57.4
Eight or more	11.8	13.9	22.9	6.8	15.4	21.3
Estimation of popularity with other students						
Very popular or popular	29.4	18.7	14.6	26.7	16.9	10.6
Average	65.6	76.5	79.1	56.6	76.9	73.6
Unpopular or very unpopular	0.0	2.4	4.2	0.0	3.1	15.8
No answer	5.0	2.4	2.1	16.7	3.1	0.0
N	17	166	48	30	225	19

number precluded any analysis of this category. Those indicating church groups (42 cases), work groups (16 cases), and social clubs (4 cases) were combined into a single category hereafter labelled "organized groups."

FACTORS ASSOCIATED WITH CHOICE OF REFERENCE GROUP

Table 1 indicates some background characteristics of the boys and girls, as they report them themselves, which might be thought to direct the selection of the different reference groups. These background traits probably existed before the reference group was chosen and hence could have influenced the choice. There is no consistent association of the degree of life satisfaction or happiness of the parents and the tendency to choose the immediate family as the reference group. Only a slightly larger percentage of boys who indicated "immediate family" as their reference group said their fathers were very happy, as compared to those who chose other reference groups, but these same boys reported their mothers to be slightly less happy, on the average. Among the girls, even these differences did not appear. If happiness of parents does not help to direct the choice of reference group, it might be thought that this important background variable is blunted in its influence on children possibly because it stems from, and reflects itself in, activities of the parents outside the home rather than within it. But data on the parents' participation in voluntary associations also show no consistent relationship to the children's choice of reference group.

The number of siblings, however, does show a relationship to the choice of reference group: In families where there are few or no siblings, children are more likely to choose school chums as their reference groups, whereas in large families, children are more likely to choose organized groups as their reference groups.

Concerning the child's estimation of his popularity with his schoolmates, it is difficult to say whether this is possible cause or possible effect of his choice of reference group.. There is a significant association between these two, however: Children who choose school chums as a reference group are more likely than other children to consider themselves popular. On the other hand, children who choose organized groups as their reference groups are least likely to think of themselves as popular with their schoolmates. These relationships are found among both boys and girls and are statistically significant at the 90 per cent level of confidence, which criterion of significance is used throughout this paper.

THE INFLUENCE OF THE REFERENCE GROUP ON SOCIAL PARTICIPATION AND CAREER PLANS

We turn now to some of the expected *consequences,* in behavior and attitude, of choice of certain reference groups by our rural high school

TABLE 2

ATTITUDES AND BEHAVIOR ASSOCIATED WITH REFERENCE GROUPS

	PERCENTAGE GIVING INDICATED ANSWER AMONG:					
	Boys for whom following groups are most important			Girls for whom following groups are most important		
	School Chums	*Immediate Family*	*Organized Groups*	*School Chums*	*Immediate Family*	*Organized Groups*
Number of organized activities youth participating in*						
None	11.8	4.2	6.2	3.3	12.0	0.0
One	0.0	8.4	18.7	13.3	8.4	15.8
Two	29.4	13.3	14.6	6.7	10.2	10.5
Three or four	17.6	23.5	29.3	26.7	29.9	36.8
Five or six	23.5	29.0	14.6	30.0	21.3	26.3
Seven or eight	11.8	10.2	10.4	6.7	9.3	5.3
Nine or more	5.9	11.4	6.2	13.3	8.9	5.3
Desire to quit school now						
Yes	11.8	6.6	6.2	3.3	4.0	5.3
No	88.2	82.6	87.6	93.4	88.9	84.2
Undecided	0.0	10.8	6.2	3.3	7.1	10.5
Expect to do when finish high school†						
Go farming	35.3	24.7	27.1	3.3	1.8	0.0
Get full-time job	29.4	33.1	25.0	30.0	39.6	47.4
Vacation 3-4 mos. & then get full-time job	11.8	8.4	6.2	0.0	4.0	5.3
Get part-time paid job	17.6	10.2	8.3	23.3	14.7	15.8
Vacation 3-4 mos. & then get part-time job	0.6	1.2	2.1	0.0	2.7	0.0
Go to college or vocational school	17.6	25.3	25.0	36.7	35.6	26.3
Misc: Housework, loaf, armed serv.	0.0	12.7	14.6	6.7	12.4	5.3
Don't know	0.0	0.6	2.1	3.3	0.0	0.0
N	17	166	48	30	225	19

* The list does not include church affiliation, which explains why a few of the boys could indicate that they participated in no organized activities and yet say that their reference group was an organized group: these boys indicated more specifically that their reference group was the church.

† Percentages add up to more than 100 as some respondents gave more than one answer.

youth. There seems to be no reliable and consistent pattern of differences among the youth who have chosen the three different kinds of reference group in regard to two important matters raised in our survey (Table 2): (1) the number of organized activities the youth participates in, both in and out of school, and (2) the desire to quit school (at the time of the survey).

There is some difference, however, in regard to plans for after high school graduation. Among the boys, a slightly larger proportion expected to go into farming, or get a part-time job, or vacation for several months and then get a full-time job, if their reference group was "school chums." If their reference group was their immediate family, however, they were more likely to be planning to get a full-time paying job immediately or to go to college. If their reference group was an organized group, they were also more likely to be planning to go to college or into the armed services or to loaf.

Among the girls whose reference group was their school friends, a disproportionately large number expected to get a part-time job or go to college. Girls whose reference group was some organized group were distinctive in their tendency to expect to get a full-time paying job and not go to college.

FACTORS ASSOCIATED WITH INTIMACY OF FAMILY LIFE

By far the most frequently mentioned reference group is the immediate family, and because the degree of intimacy with the immediate family seems so important, it is analyzed further. Tables 3 and 4 present cross-tabulations with answers to the question "How close (intimate) a family life do you have with your parents and brothers and sisters?" There is a very high correlation, as might be expected, between intimacy of family life and fondness for the parents, although the correlation is not quite as high as was found to prevail among a sample of University of Minnesota students (3). The correlation is higher among boys than among girls, especially in regard to fondness for father, in spite of the fact that more girls than boys claim to be very fond of their parents.

Both a very large number and a very small number of siblings seem to be associated with intimacy of family life: The boys who have between three and seven siblings, and the girls who have between two and five siblings, report having the least intimate family life. While few students consider themselves unpopular with their schoolmates, the lowest relative popularity was found among those who have the least intimate family life. We have already seen that these are the youth whose reference group tends to be an organized group, especially a church group.

Intimacy of family life is also negatively correlated with definiteness of selection of a future occupation: Especially among the boys with the

TABLE 3

CONDITIONS ASSOCIATED WITH INTIMACY OF FAMILY LIFE

	PERCENTAGE GIVING INDICATED ANSWER AMONG:					
	Boys who say their family life is			Girls who say their family life is		
	Very Close	*Close*	*Not Very Close or Not Close at All*	*Very Close*	*Close*	*Not Very Close or Not Close at All*
Indicated fondness for father						
Very fond	87.4	54.5	22.2	82.1	56.7	44.4
Moderately fond	9.9	32.2	38.9	9.7	36.9	55.6
No particular feeling either way	0.9	7.0	27.8	0.7	2.8	0.0
Rather dislike him	0.9	0.0	5.6	0.0	0.0	0.0
No answer	0.9	6.3	5.5	7.5	3.6	0.0
Indicated fondness for mother						
Very fond	91.9	69.2	38.9	94.5	73.8	50.0
Moderately fond	7.2	24.5	38.9	2.8	22.7	44.4
No particular feeling either way	0.9	3.5	16.7	0.7	2.1	5.6
Rather dislike her	0.0	0.0	0.0	0.7	0.0	0.0
No answer	0.0	2.8	5.5	1.3	1.4	0.0
Number of siblings						
None or one	12.6	11.2	11.1	18.6	12.7	11.1
Two	20.7	11.9	11.1	10.3	12.1	16.7
Three to five	34.2	44.1	44.4	38.7	44.8	44.4
Six or seven	15.3	14.7	16.7	18.0	16.3	0.0
Eight or more	17.2	17.5	16.7	13.7	14.1	27.8
No answer	0.0	0.6	0.0	0.7	0.0	0.0
Indicated popularity with classmates						
Very popular	3.6	0.7	0.0	5.5	4.3	0.0
Popular	24.3	14.7	5.6	16.6	11.3	5.6
Average	64.9	81.8	83.3	71.0	77.4	94.4
Unpopular or very unpopular	1.8	2.1	11.1	4.9	3.5	0.0
No answer	5.4	0.7	0.0	2.0	3.5	0.0
Training for certain (specified) occupation						
Yes	50.5	55.2	72.2	57.9	62.4	66.6
No	26.1	23.1	0.0	19.3	14.2	5.6
Not training for occupation	20.7	19.6	11.1	20.7	21.3	22.2
No answer	2.7	2.1	16.7	2.1	2.1	5.6
N	111	143	18	145	141	18

Table 4

EXPECTATIONS FOR FUTURE ASSOCIATED WITH INTIMACY OF FAMILY LIFE

	PERCENTAGE GIVING INDICATED ANSWER AMONG:					
	Boys who say their family life is			Girls who say their family life is		
Expectations	*Very Close*	*Close*	*Not Very Close or Not Close at All*	*Very Close*	*Close*	*Not Very Close or Not Close at All*
For activity after high school graduation						
Farming	23.4	28.0	33.3	2.1	2.8	5.6
Full-time paid job	32.4	30.1	33.3	40.0	42.6	27.8
Vacation then full-time job	7.2	9.8	5.6	4.8	4.3	11.1
Part-time paid job	9.9	8.4	11.1	16.6	11.3	16.7
Vacation then part-time paid job	0.9	2.1	0.0	2.1	3.5	0.0
College	34.2	16.1	16.7	29.0	35.0	38.9
Misc: Housework, loaf, armed serv.	12.6	14.7	27.8	11.7	12.8	16.7
Don't know	0.9	0.7	0.0	0.7	0.0	0.0
For staying or leaving community						
Stay	42.3	39.9	33.3	31.7	31.9	22.2
Leave	46.9	46.8	61.1	62.1	61.0	72.2
Don't know	8.1	12.6	5.6	5.5	5.0	5.6
No answer	2.7	0.7	0.0	0.7	2.1	0.0
For age of marriage (in years)						
Under 20	0.0	1.4	0.0	7.6	7.8	11.2
20–21.9	13.5	9.8	33.3	37.2	40.4	27.8
22–24.9	37.9	29.3	33.3	32.4	35.5	27.8
25–27.9	23.4	25.9	16.7	6.9	5.7	11.1
28 or over	3.6	8.4	5.6	4.2	2.1	11.1
Don't want to marry	4.5	5.6	11.1	2.1	2.1	5.5
Don't know or no answer	17.1	19.6	0.0	9.6	6.4	5.5
For future number of children (wanted)						
None	4.5	9.1	22.2	4.1	6.4	5.6
One or two	17.1	24.5	33.3	22.1	20.6	16.7
Three to five	44.2	40.6	22.3	51.7	53.2	55.6
Six to eight	8.1	4.2	0.0	9.0	9.2	16.7
Nine or more	9.9	4.2	16.7	6.2	2.1	0.0
Don't know or no answer	16.2	17.4	5.5	6.9	8.5	5.4
N	111	143	18	145	141	18

most intimate family life there are the largest proportion who are not training for any definite occupation. (There is no status differential in the occupations toward which the youth in the three categories of family intimacy are aiming, except that a larger proportion of boys with the least intimate family life are training to be farmers.)

This greater definiteness of choice of occupation among those with the loosest family ties reflects itself in the youngsters' plans for the period immediately following graduation: Those boys most intimate with their families have a greater expectation of going on to college, while those least intimate with their families are more likely than the others to expect that they will go into farming or do some unusual thing. While more boys are planning to leave the community than to stay, a larger proportion of those least intimate with their families are planning to leave and this probably has something to do with their choice of miscellaneous plans for the period following graduation.

The pattern among girls is somewhat different: While the girls with least intimate family ties also are most likely to indicate that they have miscellaneous unusual plans for the post-graduation period, they differ from the boys in this category by including the largest proportion who are planning to go on to college. A greater majority of the girls than of the boys, and again the proportion is greatest among those with least intimate family ties, are definitely planning to leave the community when they become adults.

Plans for their private life are also most unusual for those with the least intimate family life. A larger proportion of them are planning to get married either at a very young age, at a relatively advanced age, or not at all. A larger proportion of the boys among them want either no children or a very large number of children (the girls do not vary in this respect with the degree of intimacy of family life). The same tendency, to expect or plan on a relatively unusual adult life, was also found among those with least intimate family ties in the earlier study of University of Minnesota students. The evidence of the earlier study was that "responsible" behavior and "responsible" planning for adult life was most likely to be found among those with the most intimate family life, and the evidence from the present study—while not as definitive as the earlier one—suggests that this is also true among rural high school youth.

Factors Associated with Peer Group Popularity

We turn now to a closer analysis of another aspect of reference group behavior—popularity with schoolmates. We have already seen that popularity is somewhat associated with choosing school chums as the chief reference group. A much stronger correlation is noted in Table 5 between popularity and optimism regarding future life chances. Whether the estima-

TABLE 5

EXPECTATIONS FOR FUTURE ASSOCIATED WITH POPULARITY IN SCHOOL

	PERCENTAGE GIVING INDICATED ANSWER AMONG:					
Expectations	*Very Popular Boys and Girls*	*Popular Boys*	*Average Boys*	*Popular Girls*	*Average Girls*	*Unpopular Boys and Girls*
For future chances in life						
Very good chance	50.0	22.4	15.4	36.6	17.5	15.0
Good chance	22.2	57.2	49.6	43.9	41.0	25.0
Average	16.7	14.3	32.2	17.1	36.7	50.0
Poor chance	0.0	2.0	1.4	2.4	3.5	10.0
Very poor chance	11.1	0.0	0.0	0.0	0.4	0.0
Don't know or no answer	0.0	4.1	1.4	0.0	0.9	0.0
For age of marriage (in years)						
Under 20	16.7	0.0	1.0	5.8	6.9	5.0
20–21.9	50.0	16.3	11.5	37.0	37.1	20.0
22–24.9	0.0	46.9	31.2	38.1	37.1	25.0
25–27.9	5.6	20.4	26.0	8.0	7.9	10.0
28 and over	11.1	2.0	6.2	2.9	3.1	10.0
Don't want to marry	0.0	4.1	6.7	1.6	1.7	15.0
Don't know or no answer	16.6	10.3	17.4	6.6	6.2	15.0
For number of children (wanted)						
None	0.0	6.1	8.2	4.9	4.8	25.0
One or two	27.8	20.4	22.5	17.0	21.8	10.0
Three to five	33.3	51.0	39.9	63.5	53.4	30.0
Six to eight	5.6	6.1	5.8	4.8	10.9	5.0
Nine or more	22.2	2.0	7.7	2.4	2.6	15.0
Don't know or no answer	11.1	14.4	15.9	7.4	6.5	15.0
For staying in or leaving community						
Stay	27.8	32.7	41.8	26.8	31.4	45.0
Leave	38.9	57.1	46.1	68.3	62.4	55.0
Don't know	33.3	8.2	10.6	4.9	4.4	0.0
No answer	0.0	2.0	1.5	0.0	1.8	0.0
N	18	49	208	41	229	20

tion of personal popularity is based on the realities of the situation or not, apparently it and optimism reflect an important aspect of the youth's outlook on life. A curious finding here is that the proportion planning to stay in the community increases with progressively *lower* popularity. While this is partly a function of the concentration in the very popular group of those who say they do not know their future plans concerning staying in the community, it also seems to reflect important differences in general outlook on life.

Expectations for getting married at either a very young or a very late age do not seem to be correlated with estimation of popularity. But there is a significant minority among the unpopular who say they do not want to get married at all, which is not found among the other groups of youngsters. A similar pattern is revealed in response to the question concerning the number of children wanted. A fourth of the unpopular group do not want any children at all, which is a much higher percentage than that found in any other popularity group. Those wanting a very large number of children are concentrated in the two extreme groups in terms of popularity.

Reference Group Values and Expectations for the Future

Expectations for future adult life are linked to reference groups at least partly through the values which these reference groups hold. We sought to get at the relevant values of two major reference groups—parents and school chums—as well as personal aspirations for the future, by means of three projective-type questions:

1. Name three qualities or achievements you might have when you become an adult for which you think your parents would praise you.
2. Name three qualities or achievements you might have when you become an adult for which you think your school friends would praise you.
3. Name three qualities or achievements that you yourself would like to have as an adult.

Answers evoked by these directions are presented in Table 6, cross-tabulated by the reference groups of the boys and girls. The most general finding, which is supported by the percentages at several points in the table, is that where there is a discrepancy between what the parents and school friends would praise the subject for, those subjects whose reference group is their family tend to have personal aspirations reflecting the parents' values, while those whose reference group is their "school chums" tend to have personal aspirations reflecting the school friends' values. Other noteworthy conclusions that can be drawn from the table are:

1. The most frequently mentioned aspiration is an occupation. It is significantly more frequently mentioned by those whose reference group is "school chums" than by those whose reference group is "immediate family" or "organized groups." This differential is maintained when the occupation is farming. This orientation toward occupation on the part of those whose reference group is school chums—which seems to be encouraged both by their parents and by their school chums—helps to explain the previously mentioned finding that individuals in this category are the ones most likely to be planning to go to work when they graduate from high school.

2. Material success is also very frequently mentioned as an aspiration. In most cases when success is the aim, occupation is specified. Boys are more likely than girls to mention occupational success as an aspiration for

TABLE 6

REFERENCE GROUPS AND VALUED QUALITIES OR ACHIEVEMENTS

(percentages naming indicated qualities or achievements)*

	Qualities or achievements praised BY PARENTS						Qualities or achievements praised BY SCHOOL FRIENDS						Qualities or achievements desired BY SELF					
	Among *boys* whose reference group is:			Among *girls* whose reference group is:			Among *boys* whose reference group is:			Among *girls* whose reference group is:			Among *boys* whose reference group is:			Among *girls* whose reference group is:		
	School Chums	*Family*	*Organized Groups*	*School Chums*	*Family*	*Organized Groups*	*School Chums*	*Family*	*Organized Groups*	*School Chums*	*Family*	*Organized Groups*	*School Chums*	*Family*	*Organized Groups*	*School Chums*	*Family*	*Organized Groups*
1. Get a job or specific occupation (excluding farming)	68	28	45	100	57	69	73	21	25	85	48	49	97	20	38	84	45	42
2. Be a farmer	39	17	18	5	0	0	36	13	16	0	1	0	35	16	19	0	0	0
3. Be successful: Generally	39	29	27	10	20	17	9	26	13	20	19	35	0	25	25	20	18	12
In occupation	48	50	33	19	38	23	55	43	32	15	29	28	62	57	44	15	31	42
In style of life	10	15	3	5	5	6	9	17	10	10	4	0	9	24	6	10	9	12
4. Be a good family member	0	7	3	38	17	6	0	0	3	20	9	21	0	2	0	20	12	30
5. Have a good family and home life	19	21	18	5	19	29	18	15	16	25	19	21	9	35	28	25	33	60
6. Be socially adjusted and popular	0	12	12	29	23	23	0	46	28	35	43	21	9	18	19	34	39	30
7. Have a nice appearance	0	4	0	0	1	0	0	2	0	5	4	7	0	2	9	5	3	0
8. Have good character traits	10	43	21	33	37	23	18	38	38	35	40	21	27	27	9	34	32	12
9. Be religious	0	13	15	19	13	0	0	5	10	10	7	7	0	4	9	10	10	6
10. Be a good citizen and help others	0	6	9	5	9	0	0	10	13	0	11	0	0	7	13	0	11	0
11. Have talent: Mental	19	20	27	19	26	17	0	11	19	10	18	28	0	14	9	10	16	6
Physical	10	9	15	10	11	6	9	15	16	10	13	7	18	11	13	10	9	0
12. Be happy and do what I like	10	5	9	0	10	23	36	12	16	15	4	21	27	21	22	15	19	24
13. Miscellaneous and trivial	10	1	0	0	0	0	0	1	0	0	2	0	0	1	0	0	1	0
14. Don't know	18	20	45	3	14	58	37	25	45	5	29	34	7	16	37	8	12	24
Total	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300
N	17	166	48	30	225	19	17	166	48	30	225	19	17	166	48	30	225	19

* Percentages add to 300 per cent as each subject was asked to give three responses to each question. Cases of "no answer" were excluded from the calculations.

themselves and as a future achievement for which their parents and school friends would praise them.

3. Social adjustment and popularity are more frequently mentioned as personal aspirations and as achievements which school friends would praise than as achievements which parents would praise. In general girls think more of these traits than do boys.

4. Girls are also more likely than boys to think of being a good family member. They are especially likely to consider this a goal for themselves if their reference group is an organized group rather than their family or school chums. But when guessing what their parents would praise them for, girls are more likely to mention being a good family member when their reference group is their school chums.

5. Character traits and talents which are the goals of the ideal typical "inner-directed" personality are less frequently mentioned as aspirations and praised goals than are material success and occupations that we have already considered. This fact supports Riesman's hypothesis that other-directed characteristics are more highly valued than inner-directed ones (2). Notably, being religious and being a good citizen are infrequently mentioned. Among boys, those whose reference groups are organized groups are more likely to mention these latter goals than those whose reference groups are family or school chums.

6. Good character traits are generally as likely to be mentioned as personal aspirations as they are to be considered as something which parents and school friends would praise. This fact contradict's Riesman's hypothesis that the *trend* is away from evaluating good character highly, at least as perceived by the youngsters in our study.

7. Those whose reference groups are organized groups are consistently most likely to say they "don't know" what their personal aspirations are and what their parents and school friends would praise them for. It is this category of our sample which is thus least oriented toward their future as adults.

8. In addition to Riesman's categories of "inner-directed" and "other-directed,"[2] one might distinguish another broad set of life-goals which is represented by our category 12, "Be happy and do what I like." We might label this the "gratification-directed" or hedonic personality. It approximates the ideal formulated by psychoanalytic theory. That it may be an increasingly frequent type of personality is suggested by the fact that it is mentioned more frequently as a personal aspiration and as a trait which school friends would praise than one which parents would praise. Further, those whose reference group is the immediate family (that is, those particularly oriented toward their parents) are less likely to refer to this trait than are those

---

[2] Riesman has a third category of "tradition-directed" which we have not been able operationally to distinguish from "inner-directed" in this study.

whose reference groups are school chums (for the boys) or organized groups (for the girls).

SUMMARY

This study has examined just a few of the background characteristics, behaviors, and attitudes that might be hypothesized to be associated with the different membership reference groups held among rural high school youth. The statistically significant differences among youth with different reference groups were in regard to number of siblings, estimation of popularity among schoolmates, and plans for the period after graduation.

The most frequently mentioned reference group was the immediate family, and analysis was made of various characteristics associated with different degrees of reported intimacy of family life. Intimacy of family life was very highly correlated with fondness for father and mother, and was more characteristic of children who had either very few or very many siblings. Those very close to their families were also more likely than others to be popular with their schoolmates, uncertain about their future vocation, planning to leave the community when adult, get married at an average age and have a moderate number of children. Planning to go to college showed different relationships by sex: Among boys, it was those who are closest to their families who are planning for college, while among girls the opposite was true.

Popularity was associated with the desire to leave the community when adult and with optimism regarding future life chances. Those who did not want to get married or have children were concentrated among the unpopular.

Aspirations regarding personal qualities or achievements were found to be related to what were perceived to be the values of one's reference groups. Evidence was found supporting Riesman's hypothesis that qualities of the "other-directed personality" are more highly valued than are qualities of the "inner-directed personality." But there was no discrepancy in this regard between the youths' own values and what they perceived their parents' values to be. There was a tendency, however, for the youth to aspire to qualities of what might be called the "gratification-directed" or "hedonic" personality more than their parents were perceived to do so.

REFERENCES

1. HYMAN, HERBERT. The psychology of status. *Arch. Psychol.*, 1942, No. 269.
2. RIESMAN, DAVID. *The lonely crowd.* New Haven: Yale Univer. Press, 1950.
3. ROSE, ARNOLD M. Acceptance of adult roles and separation from family. Unpublished manuscript.

FILM-MEDIATED FANTASY AGGRESSION AND STRENGTH OF AGGRESSIVE DRIVE[1]

ALBERTA ENGVALL SIEGEL
The Pennsylvania State University

Those concerned with healthy personality development have shown considerable interest in the effects which the mass media may have upon children. Much of this interest has centered on the possible harmful effects of the aggressive and hostile themes which predominate in many entertainment films, commercial television programs, and comic books.

However, one frequently encounters the reassuring assertion that children's involvement in such themes may provide a useful "outlet" for their aggressive impulses. This "release" notion is expressed in psychoanalytic theory in the concept of catharsis, has been codified in behavior theory, and is an accepted and intrinsic part of many theories of psychotherapy. Its usefulness in an evaluation of the effects of violence in the mass media has been widely accepted, but we lack experimental evidence which either affirms or infirms its worth. The present study is one attempt to subject to experimental test the "release" theory as it applies to the effects of violence in the mass media.

THE HYPOTHESIS OF EQUIVALENCE OF FORMS

Behavior theorists have given considerable attention to the dynamics of aggressive motivation. They have reasoned that aggression develops as a secondary drive early in the socialization process (21, 22), and that in

---

[1] Professors Lois Meek Stolz, Quinn McNemar, Paul R. Farnsworth, and Robert R. Sears, all of Stanford University, were most helpful to the writer in the planning of this study. Professors C. R. Carpenter, Winona L. Morgan, and Robert G. Bernreuter gave generous administrative support to the execution of the study, which was conducted while the writer was a research fellow of the Instructional Film Research Program at The Pennsylvania State University, supported by a grant from the Carnegie Corporation of New York. The writer is most grateful to Dr. Ellen Tessman for her unselfish and highly skilled service as the observer in this study. The help of Leslie P. Greenhill and other staff members of the Instructional Film Research Program, and of Jane A. Bovie, Mary C. Lane, and Barbara W. Norman, staff members in the Department of Child Development and Family Relations at The Pennsylvania State University, is gratefully acknowledged.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

the person's interaction with his environment it occurs as a response to frustration (2). Because aggressive behavior so frequently evokes retaliation or punishment, direct acts of aggression may come to be replaced by more indirect acts (be attenuated) or to be deflected to less threatening objects or persons (be displaced). Thus many modes of expressing aggression are developed. Fantasy aggression is one of these modes (2, p. 10, p. 46).

Behavior theorists reason that these alternative ways of expressing aggression are dynamically related in that each may serve to reduce aggressive drive. The assertion that any of a diverse variety of aggressive acts can reduce aggressive drive is *the hypothesis of equivalence of forms,* the behavior theory formulation of the "release" notion: "the expression of and act of aggression is a catharsis that reduces the instigation to all other acts of aggression" (2, p. 53).

Evidence for the phenomenon of displacement of aggression has repeatedly been secured (7, 10, 14, 15, 16, 24, 29, pp. 276-280, 31). Moreover, various investigators have found systematic relations between fantasy and overt aggression (6, 17, 18, 19, 20, 27). These findings provide affirmative evidence for necessary preconditions to the antecedent-consequent relation asserted by the hypothesis of equivalence of forms, but none bears directly on that hypothesis.

Direct confirmation for that hypothesis was given by Lewin *et al.* (9) in certain observations that scapegoating by children is followed by a reduction in over-all level of aggression. Stone (26), in an unpublished study reported by McClelland, analyzed the relations among various forms of aggression in college football players. His findings "tend to throw some doubt on the idea that there is a kind of reservoir of aggression which may be lowered by expression through various channels which may substitute for each other if one happens to be blocked" (11, p. 517). However, direct confirmation for the hypothesis as it applies to fantasy aggression has been obtained by Feshbach (3) in his ingenious study of the effect of the experience of writing aggressive TAT fantasies upon situationally-aroused aggression.

The present study tests two hypotheses which were derived from the hypothesis of equivalence of forms. Both concern the effect of fantasy aggression upon subsequent level of aggressive drive. Since aggressive drive is an intervening variable, not subject to direct measure (22, p. 227), two indices of aggressive drive are used: level of aggressive behavior and level of anxiety and guilt. Anxiety is taken as an index to aggressive drive because it is reasoned that in the socialized middle-class child the activation of aggressive drive directly arouses anticipation of punishment (either by others or by the self), which is anxiety. If anxiety accompanies the activation of drive, then changes in drive strength should be accompanied by changes in level of anxiety. The use of level of anxiety as well as level of aggressive behavior as separate indices to drive strength seemed necessary because of the complex relation which may well exist between aggressive drive strength and overt social aggression in the socialized child.

The two research hypotheses were:

I. *Fantasy aggression reduces the instigation to all other acts of aggression. With the level of aggressive drive held constant, there should be an inverse relation between the expression of fantasy aggression and the subsequent expression of other forms of aggression.*

II. *Fantasy aggression reduces the institgation to all other acts of aggression and therefore reduces aggression anxiety, since the latter is presumed to be a direct correlate of the former. With the level of aggressive drive held constant, there should be an inverse relation between the expression of fantasy aggression and the subsequent behavioral expression of anxiety.*

Procedure

The hypotheses were tested by comparing the play behavior of children after a film-mediated fantasy aggression experience with their play behavior after seeing a control (nonaggressive) film. Each film was presented at a different session, and sessions were separated by a period of one week. Aggressive drive was "held constant" by using each child as his own control. Each child attended both sessions with the same partner, a child of his own sex from his nursery school group. After each film showing, the two children were left alone in a playroom for a 14-minute period, and their play during that time was scored for aggression and for anxiety and guilt by scorers behind a one-way vision mirror.

Subjects

Twenty-four children, 12 boys and 12 girls, participated in this study. Their ages ranged from 3-9 to 5-1. All were enrolled in nursery schools. The children were assigned to partnerships by their nursery school teachers, who were asked to constitute each pair so that the two children would be willing to go with each other and the experimenter to an unfamiliar building for a new experience.

Procedures with the Children

Participation in the study was secured by inviting the scheduled children "to see a movie." The pairs of children accompanied the experimenter to a playroom in another building from that in which their school was housed. The film projector in that room had been set up in advance, and the film was begun as soon as the children were seated in the room. The experimenter observed and rated the children's behavior while the film progressed. When the film ended, she told the children that she must leave "to do some work in another room for awhile," and that they were to play with the toys in the room during her absence. She stressed that when she returned she would knock on the door before re-entering the room. When this information seemed to be understood by both children, the experimenter wheeled the projector from the room and left the children alone together.

The arrangement of the playroom was the same for each of the 24 sessions. These toys were available to the children: two rubber daggers, two lumps of soft clay, two toy telephones, four soft rubber sponges, a plastic tea set and miniature flatware, seven small toy vehicles, a doll bed containing a doll and bedding, eight inflated balloons, and a large inflated plastic punching toy which stood child-height.

At the conclusion of 14 minutes of scoring of the children's play by the observer and the experimenter, while observing through the one-way mirror, the experimenter returned to the room, terminated the play, and accompanied the children back to their nursery school.

Scoring Procedures

During each film showing, the experimenter assigned two ratings to each child: one on interest in the film, and one on behavioral signs of anxiety. The five-point interest scale[2] was adapted from that developed by Dawson (1) for use in evaluating the entertainment value of children's films. The four-point anxiety scale was developed for this study and made use of a 29-item list of behavioral indicators of anxiety (including such signs as these: grips chair, winces, swallows, wets lips, sucks thumb, breathes rapidly, etc.).

At the conclusion of the film showing, the observer was summoned by a third person to the observation room. She waited in a sound-proof room in another part of the building during each film showing. This isolative procedure prevented the observer's knowing which film had been shown to any children, and thus precluded the possibility that her scoring of the play might be biased with respect to the hypotheses under test.

The observer was joined by the experimenter in the observation room, and both scored the play of both children during a 14-minute period of free play. Their technique of scoring aggression and anxiety and guilt is discussed below. The scores assigned by the observer during the play periods are the ones used in all data analyses; those of the experimenter were used to check the reliability of the observer's scoring.

During the weeks in which the experiment was conducted, the children's nursery school teachers were asked to rate the children on their habits of aggression in the nursery school, using slight modifications of scales developed by Gewirtz (4). These ratings were assigned to 18 of the 24 subjects (the other six being members of other nursery school groups).

---

[2] Complete copies of the scales used in this study are given in (23).

Experimental Controls

In order to isolate the effects of the variables under study, certain controls were instituted. Some of these have already been mentioned, but all are summarized below:

1. Any child's partner in play was the same person for both play sessions. Thus any differences between sessions could not be attributed to systematic differences in interpersonal stimulation.

2. Sequence effects were randomized. That is, half of the pairs saw the aggression film at their first session and the control film at their second, whereas the other pairs saw the films in the reverse order. Thus any differences between the E and C play sessions could not be attributed to systematic (or sequential) differences between sessions regardless of films.

3. Effects of age, previous experience, and social background were controlled by using each child as his own control. That is, comparisons between each child's behavior under the two conditions constitute the data of the study. This design also serves to "hold aggressive drive constant," as is required by the hypotheses under test.

4. Timing effects were controlled by separating the E and C sessions by an interval of time which was the same for all pairs. The effort to separate the two sessions by just seven days was successful in all but two cases: for 10 pairs the two sessions were separated by seven days, for one pair the two sessions were separated by six days, and for one pair the two sessions were separated by 11 days.

5. Time-of-day effects on level of aggression were held constant; any pair participated in both E and C sessions at the same time of day.

6. Observer bias was controlled by preventing the observer from knowing whether any session was an E or C session. The scores of the observer were used in all data analyses.

7. The combination of sex and session effects was randomized. That is, an equal number of pairs of girls and boys experienced the E session first; a comparable equal number of boys and girls experienced the C session first. Three pairs of boys and three of girls had sessions in the E-then-C sequence; three other pairs of boys and three other of girls had sessions in the C-then-E sequence.

8. Insofar as possible, the two films were matched on all variables except that under test. Realistically, such matching is possible in only a general way unless films are especially produced for experimental purposes, which was not the case in this study.

OPERATIONAL DEFINITIONS

This is a study of the relation between fantasy aggression, the antecedent variable, and strength of aggressive drive, the consequent variable. The operational meanings which were attached to these terms are given in this section.

Fantasy Aggression

Fantasy aggression was mediated by a 10-minute cartoon film which was selected by three psychologists for its direct, unabashed, and easily comprehensible portrayal of extreme interpersonal aggression. This colored animated film depicts the conflicts which occur between Woody Woodpecker, a well-known picaresque character, and a large Air Force sergeant. Raw aggression and unrelenting hostility dominate almost every scene of this, the E film.[3]

The control film was an animated cartoon portrayal of the well-known fable of The Little Red Hen. This C film,[3] selected by the same judges, was matched in length and in use of color with the E film. Designed to appeal to young children, it is narrated by a calm-voiced woman.

These films were originally chosen because they presented a vivid contrast in the use and non-use of aggressive themes and because they seemed suitable to the level of understanding of young children. Empirical confirmation of the appropriateness of these films was provided by two sets of data: data on the children's interest in the films and data on their anxiety during the films.

Interest in films. Each child's interest in each film was rated[4] by the experimenter during each film showing in the course of the study. A difference score was then obtained for each child, representing the difference between his rated interest in the E film and his rated interest in the C film. A statistical test on these data is impracticable, inasmuch as it is the null hypothesis (no difference in interest between the two films) which we wish to establish. Descriptive data suggest that the films were equated in interest for these children: the mean rating of interest during the E film showings was 3.50, whereas the comparable mean for the C film showings was 3.25. The difference between these two means is insignificant. Fourteen of the 24 children received the same interest ratings for both films. Nine others received ratings which differed by only one point. Only one child's difference score was larger than one point.

---

[3] E film: Woody Woodpecker in "Ace in the Hole." A Walt Lantz "Cartune" in Technicolor. Universal Picture, MCMXLII. C film: "The Little Red Hen: Background for Reading Expression." David A. Smart. Coronet Instructional Film, MCML. The instructional and hortatory sequences in this film were cut from the print used in this research.

[4] The reliability of the ratings of interest was not determined, since the experimenter was alone in making these ratings during the study. However, during the period of the preliminary pilot study, the observer as well as the experimenter made interest ratings with eight children. For seven of these, the ratings of observer and experimenter were identical; for the other child, the two ratings diverged by but one point.

Anxiety during film. Were the films effectively different with respect to inducing fantasy aggression? The most relevant evidence for answering this question comes from the ratings of each child's anxiety during each film showing. If the E film's effect was really to produce aggressive fantasy in the child, then the child's anxiety should be greater during the E film than during the C film, since aggression is presumed to be anxiety-arousing in socialized children.

Of the 24, 16 children received a higher rating during the E film showing than during the C film showing, eight received the same rating for anxiety during both films, and no child received a higher rating for anxiety during the C film. These data were analyzed by the sign test (12, p. 357). In accordance with the usual practice, those cases whose difference scores had no sign (i.e., those eight cases whose difference scores were zero) were dropped from the analysis. For the 16 cases whose difference score did have a sign, the difference in rated anxiety between the E and C film showings is significant at the $p < .00002$ level. The finding that the children typically were more anxious during the E than during the C film showing is supporting evidence for the presumption that the E film elicted fantasy aggression. (The writer's confidence in this finding would be greater if the conditions of the rating had more closely approximated objectivity. The ratings of anxiety during the film showings are deficient in two important respects: (*a*) their reliability is not known, and (*b*) they were made by the experimenter with full knowledge of which film the children were watching.)

Strength of Aggressive Drive

During the play periods after each film showing, the play of both children was scored by the observer and the experimenter for level of aggression and level of anxiety and guilt.

The scorers each entered four scores at every 20-second interval, paced by an electric interval timer. Each child received two scores from each of the two independent scorers: one for level of aggression in play and one for level of anxiety and guilt. The scores for each 20-second interval of play ranged from 0 to 3.

Aggression in play. Each scorer observed the play of each child for behavior which might be termed "hostile," "destructive," or "aggressive." Such behavior might be toward the other child, toward a toy, or toward the self. Scorers were instructed to judge the intensity of any aggressive act in terms of both the quality of the instrumental act and the nature of the goal response. If no aggression was observed during the 20-second interval, a 0 was entered. If aggression was observed, its intensity was represented by scores of 1, 2, or 3.

In the course of 14 minutes, 42 separate aggression scores were given to each child. These 42 were summed to obtain the child's total aggression score for the play session. These totals (from the observer's protocols) were used in all data analyses.

The agreement between the observer and the experimenter in scoring was tested by comparing the totals from the observer's protocols with those from the experimenter's protocols. Inspection revealed that a linear relation obtained between these two sets of scores, and their agreement (i.e., the reliability of the aggression scores) is expressed by the Pearson correlation coefficient, $r = .97$.

This coefficient shows that the two scorers' data on aggression were in unusually high agreement. The high inter-scorer reliability is probably attributable to three factors: (*a*) the intensive self-training of the observer and the experimenter during the pilot study period which preceded the conduct of the study being reported, (*b*) the sensitivity of the scoring technique—the total scores depend on many specific judgments made during very short periods of play, and (*c*) the fact that the sort of aggression which young children display in this free play situation in the absence of adults is much more open, direct, and extreme than that usually seen by students of child behavior.

Anxiety and guilt during play. Each scorer observed the play of each child for behavioral signs of anxiety and guilt. The scoring was guided by a specific list of signs, adapted from the work of Ryder (25, pp. 348-349), which suggest that the child is controlling, suppressing, repressing, inhibiting, or denying his own aggressive impulses or those of his partner. Included in the list were such behaviors as: moralistic comments; tense, stiff, or awkward movements; incompleted or blocked aggressive acts; tics, thumb-sucking, and other "nervous mannerisms"; tense mouth; mask-like expression, etc.

As with the aggression scores, a score of 0, 1, 2, or 3 was entered for each 20-second interval. The child's total score for anxiety and guilt in play was the sum of the 42 separate scores assigned to him during the play period by the observer.

Inspection revealed that a linear relation obtained between the observer's and the experimenter's scores on anxiety and guilt, and the reliability of these scores is expressed by $r = .77$.

This coefficient indicates adequate inter-scorer agreement. The lower reliability of the anxiety and guilt scores, compared to the aggression scores, seems to be a function of the difficulty of the judgment as well as of the relative infrequency of such behavior in this free play situation. That is, the infrequency of occurrence of anxious and guilty behavior meant that the scores of the group were homogeneous, a fact which tends to depress the correlation coefficient.

Operational Statements of the Hypotheses

Having given operational meaning to the terms of the two hypotheses, we may now state them in the form in which they were tested:

I. *Typically, a child's total aggression score for his free social play after the showing of the E film will be lower than his total aggression score for his free social play after the showing of the C film, under the conditions of this experiment.*

II. *Typically, a child's total score on anxiety and guilt for his free social play after the showing of the E film will be lower than his total score on anxiety and guilt for his free social play after the showing of the C film, under the conditions of this experiment.*

RESULTS

Hypothesis Tests

To test Hypothesis I, a difference score was obtained for each child, representing the numerical difference between his aggression score for the C play session and his aggression score for the E play session. These difference scores were then tested for significance by the Wilcoxon matched-pairs signed-ranks test, a nonparametric test for the significance of the difference between the central tendencies of two related sets of scores (30). This test was used in preference to the more frequently seen t test because for these data it has three important advantages over that test: (*a*) it makes no assumption of underlying normality in the population of scores from which the sample was drawn, (*b*) it permits the various matched pairs to be drawn from different populations, as was done in this study—the two populations being boys and girls, who are known to differ in aggression (28)—and (*c*) it does not assume that the scores used in the analysis are sufficiently exact to be properly subject to the operations of arithmetic. By the Wilcoxon test, $T = 110.5$, a value which is not significant. Somewhat more aggression occurred in the E than the C sessions, but the difference was not significant.

To test Hypothesis II, a similar analysis was made of the scores on anxiety and guilt in play. Somewhat more anxiety and guilt occured in the E sessions (i.e., the difference was in the opposite direction from that predicted), but by the Wilcoxon test ($T = 79$) the difference is insignificant.

Thus, neither in aggression nor in anxiety and guilt were there significant differences between the E and C play sessions, and therefore neither Hypothesis I nor Hypothesis II was confirmed. Those differences which did occur, though insignificant, were in the opposite direction from that predicted. Certain other findings may aid in the interpretation of these results; these are reported below.

Other Findings

By the Mann-Whitney U test (13), a nonparametric technique, the scores were tested for sex differences. The boys showed significantly ($p < .0002$) more aggression in play than the girls, and significantly ($p < .05$) more anxiety and guilt in play than the girls.

The 48 aggression scores were correlated with the 48 anxiety and guilt scores. For these data, $r = .30$ ($p < .05$). Inspection revealed that the relation between the two sets of scores deviates from linearity, a fact which makes the Pearson r a conservative index of the degree of correlation between the two variables. The observed correlation between anxiety and aggression gives some support to the original assumption that both sorts of behavior are dynamically related to strength of aggressive drive.

For the 18 children who came from the same nursery school, a single combined aggression score was obtained for each child by summing the child's two aggression scores (for the E and C play sessions). These combined scores were correlated with the ratings which the children's two teachers had assigned to them. The two teachers rated independently, and their ratings were summed to provide a single group of scores for this analysis. One scale asked the teachers to rate "How often does the child attack other children physically?" The inter-teacher agreement in using this scale was linear and is expressed by the Pearson correlation, $r = .74$. The correlation between these ratings and the children's combined aggression scores, representing their aggression in the two play sessions, is $rho = .68$ ($p < .002$) (8, p. 48). The other scale asked the teachers to rate "How often does the child destroy or damage the property of another child?" The inter-teacher agreement in using this second scale was also linear and is expressed by $r = .69$. The correlation between these ratings and the children's combined aggression scores is expressed by $rho = .56$ ($p < .02$).

Significant differences in aggression and in guilt and anxiety were found between Sessions I and Sessions II, irrespective of which was the E and which the C session. These differences will be reported in a separate paper.

DISCUSSION

Because of the nature of statistical inference, acceptance of the null hypothesis is a weaker conclusion than its rejection. Acceptance of the null hypothesis for a given set of data does not necessarily require rejection of the research hypothesis, or denial of its validity. Before such a strong interpretation is put on the finding of no differences, certain alternative interpretations of the results of the hypothesis tests must be considered.

In the first place, the measurement of the variables may be examined for sensitivity and appropriateness.

The evidence is that the measures of aggression and of anxiety and guilt were sufficiently sensitive. This is indicated by (*a*) the good inter-scorer agreement, (*b*) the fact that the measures discriminated between sexes, (*c*) the fact that the measures discriminated between Sessions I and Sessions II, and (*d*) the significant correlations between the aggression scores and the teachers' ratings of habits of aggression.

The appropriateness of the measures as indices of strength of aggressive drive may, however, be questioned. It seems likely that the aggression scores were more directly reflective of habit strength than of drive strength. This interpretation is supported by the finding that teachers' ratings of habits of aggression correlate significantly with the children's aggression scores for the play sessions. It is also supported by the extremely significant sex differences.

Second, we should point out that only one E and one C film were used in this study, and that it would be quite inappropriate to generalize from the findings of this study to the effect of all films (5). Moreover, perhaps the experience of fantasy aggression in young children occurs most strongly with stimuli which are immediately and intimately familiar to them. The hypothesis deserves a test with films which portray aggression by young children in a familar environment—in a home, neighborhood, or nursery school.

Finally, it may be that the application of the hypothesis of equivalence of forms to drive-produced aggression is premature. The hypothesis was originally formulated in behavior theory to apply to frustration-produced or situationally-aroused aggression, and perhaps we should seek its verification in this context before applying it to drive-produced aggression.

Before concluding this discussion, we should caution that this study was addressed to only one possible effect of film violence: the effect on strength of aggressive drive. Violent films may well have other effects, particularly on children's beliefs, their role-perceptions, and perhaps their values and attitudes. No evidence on these possible effects was gathered in this study. Further research on the effects of violence in the mass media is necessary before we can even begin to attempt a comprehensive and well-grounded appraisal.

SUMMARY

The behavior theory hypothesis of equivalence of forms was tested as it applies to the effects of film-mediated fantasy aggression on strength of aggressive drive in young children. The free play of like-sexed pairs of children in the absence of adults was observed under two conditions: after a highly aggressive cartoon film and after a "matched" nonaggressive film. Play was scored for aggression and for behavioral signs of guilt and anxiety. These scores did not differ significantly for the two conditions. Significant

sex differences and session differences were found, and the children's aggression scores correlated significantly with teachers' ratings of habits of aggression in the nursery school. These findings suggest that the acceptance of the null hypothesis with respect to the hypothesis of equivalence of forms may be due to the inappropriateness of the scores as indices of drive strength, and that the hypothesis of equivalence of forms should be tested with films of greater immediacy and in a situation in which frustration-produced aggression is aroused.

REFERENCES

1. DAWSON, MARJORIE G. *The children's film library and special children's programs.* (4th Ed.) New York: National Children's Film Library, 1949.
2. DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.
3. FESHBACH, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.,* 1955, 50, 3-11.
4. GEWIRTZ, J. L. Succorance in young children. Unpublished doctor's dissertation, State Univer. of Iowa, 1948.
5. HAMMOND, K. R. Representative vs. systematic design in clinical psychology. *Psychol. Bull.,* 1954, 51, 150-159.
6. HOLLENBERG, ELEANOR, & SPERRY, MARGARET. Some antecedents of aggression and effects of frustration in doll play. *Personality,* 1950, 1, 32-43.
7. HOVLAND, C. I., & SEARS, R. R. Minor studies of aggression: VI. Correlation of lynchings with economic indices. *J. Psychol.,* 1940, 9, 301-310.
8. KENDALL, M. G. *Rank correlation methods.* London: Griffin, 1948.
9. LEWIN, K., LIPPITT, R., & WHITE, R. K. Patterns of aggressive behavior in experimentally-created "social climates." *J. soc. Psychol.,* 1939, 10, 271-299.
10. LINDZEY, G. An experimental examination of the scapegoat theory of prejudice. *J. abnorm. soc. Psychol.,* 1950, 45, 296-309.
11. MCCLELLAND, D. C. *Personality.* New York: Dryden, 1951.
12. MCNEMAR, Q. *Psychological statistics.* (2nd Ed.) New York: Wiley, 1955.

13. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.,* 1947, 18, 50-60.

14. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.,* 1948, 43, 155-178.

15. Miller, N. E., & Bugelski, R. Minor studies of aggression: II. The influence of frustrations imposed by the in-group on attitudes expressed toward out-groups. *J. Psychol.,* 1948, 25, 437-442.

16. Murray, H. A. The effect of fear upon estimates of maliciousness of other personalities. *J. Psychol.,* 1933, 4, 310-339.

17. Murray, H. A. *Thematic apperception test manual.* Cambridge, Mass.: Harvard Univer. Press, 1943.

18. Mussen, P. H., & Naylor, H. K. The relationships between overt and fantasy aggression. *J. abnorm. soc. Psychol.,* 1954, 49, 235-240.

19. Pittluck, Patricia. The relation between aggressive fantasy and overt behavior. Unpublished doctor's dissertation, Yale Univer., 1950.

20. Sanford, R. N., *et al.* Physique, personality, and scholarship. *Monogr. Soc. Res. Child Develpm.,* 1943, 8, No. 1.

21. Sears, R. R. Frustration and aggression. In P. L. Harriman (Ed.), *Encyclopedia of psychology.* New York: Philosophical Library, 1946. Pp. 215-218.

22. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, Pauline S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.,* 1953, 47, 135-234.

23. Siegel, Alberta E. The effect of film-mediated fantasy aggression on strength of aggressive drive in young children. Unpublished doctor's dissertation, Stanford Univer., 1955.

24. Stagner, R. Studies in aggressive social attitudes: III. The role of personal and family scores. *J. soc. Psychol.,* 1944, 20, 129-140.

25. Stolz, Lois M., *et al. Father relations of war-born children.* Stanford, Calif.: Stanford Univer. Press, 1954.

26. Stone, A. A. The effect of sanctioned overt aggression on total instigation to aggressive responses. Unpublished honors thesis, Harvard Univer., 1950.

27. Symonds, P. M. *Adolescent fantasy.* New York: Columbia Univer. Press, 1949.

28. TERMAN, L. M., & TYLER, LEONA E. Psychological sex differences. In L. Carmichael (Ed.) *Manual of child psychology.* (2nd Ed.) New York: Wiley, 1954. Pp. 1115-1161.

29. WHITING, J. W. M., & CHILD, I. *Child training and personality.* New Haven: Yale Univer. Press, 1953.

30. WILCOXON, F. *Some rapid approximate statistical procedures.* Stamford, Conn.: American Cyanamid Co., 1949.

31. WINSLOW, C. N. The social behavior of cats: I. Competitive and aggressive behavior in an experimental runway situation. *J. comp. Psychol.*, 1944, 37, 297-313.

ANXIETY IN CHILDREN, SCHOOL ACHIEVEMENT, AND INTELLIGENCE

BOYD R. MCCANDLESS and ALFRED CASTANEDA[1]
Iowa Child Welfare Research Station

The authors, with Palermo (1) have previously reported on the children's form of the manifest anxiety scale (CMAS) and on some of its relationships with learning tasks (2, 3). This scale is an adaptation of the Taylor manifest anxiety scale for adults (4), which is in turn patterned after the Minnesota Multiphasic Inventory.

The purpose of this paper is to report on correlations between anxiety as defined by the score for the CMAS, academic achievement as measured by the Iowa Every Pupil Test (IEPT), and intelligence as measured by the Otis Quick Scoring Mental Ability Test, Form B (Otis). Intelligence test data are available only for the sixth grade groups of the fourth, fifth and sixth grade public school population that served as subjects for the present study.

SCHOOL ACHIEVEMENT

Table 1 presents the correlations, by grade and sex, between anxiety scores and sub- and composite scores on the IEPT. These two tests were administered within a week of each other by classroom teachers.

Thirteen of 30 computed relationships between anxiety and school achievement were found to be significant. No consistent pattern of relationship is manifested for the fifth grade population nor for fourth grade boys, but relationships run from moderate to high for fourth grade girls and sixth grade boys and girls. For groups where relationships exist, there is a tendency for the more complicated skills (reading, arithmetic and composite performance) to suffer more interference from anxiety than the simpler, mnemonic skills such as spelling. Also, where a pattern exists, there tends to be more interference for girls than for boys, with the exception

---

[1] The authors wish to express appreciation to Mr. David S. Palermo for his valuable work in helping to secure and process the data on which this paper is based; to Mr. Buford Garner, Superintendent; Mr. David Stewart, Coordinator of Elementary Curriculum; and Mrs. Sylvella Jacobsen, Psychologist, all of the Iowa City Public Schools; and the 15 Iowa City classroom teachers who administered tests, as well as the four elementary school principals concerned, for their help in making this study possible.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

28. Terman, L. M., & Tyler, Leona E. Psychological sex differences. In L. Carmichael (Ed.) *Manual of child psychology.* (2nd Ed.) New York: Wiley, 1954. Pp. 1115-1161.

29. Whiting, J. W. M., & Child, I. *Child training and personality.* New Haven: Yale Univer. Press, 1953.

30. Wilcoxon, F. *Some rapid approximate statistical procedures.* Stamford, Conn.: American Cyanamid Co., 1949.

31. Winslow, C. N. The social behavior of cats: I. Competitive and aggressive behavior in an experimental runway situation. *J. comp. Psychol.,* 1944, 37, 297-313.

ANXIETY IN CHILDREN, SCHOOL ACHIEVEMENT, AND INTELLIGENCE

BOYD R. MCCANDLESS and ALFRED CASTANEDA[1]
Iowa Child Welfare Research Station

The authors, with Palermo (1) have previously reported on the children's form of the manifest anxiety scale (CMAS) and on some of its relationships with learning tasks (2, 3). This scale is an adaptation of the Taylor manifest anxiety scale for adults (4), which is in turn patterned after the Minnesota Multiphasic Inventory.

The purpose of this paper is to report on correlations between anxiety as defined by the score for the CMAS, academic achievement as measured by the Iowa Every Pupil Test (IEPT), and intelligence as measured by the Otis Quick Scoring Mental Ability Test, Form B (Otis). Intelligence test data are available only for the sixth grade groups of the fourth, fifth and sixth grade public school population that served as subjects for the present study.

SCHOOL ACHIEVEMENT

Table 1 presents the correlations, by grade and sex, between anxiety scores and sub- and composite scores on the IEPT. These two tests were administered within a week of each other by classroom teachers.

Thirteen of 30 computed relationships between anxiety and school achievement were found to be significant. No consistent pattern of relationship is manifested for the fifth grade population nor for fourth grade boys, but relationships run from moderate to high for fourth grade girls and sixth grade boys and girls. For groups where relationships exist, there is a tendency for the more complicated skills (reading, arithmetic and composite performance) to suffer more interference from anxiety than the simpler, mnemonic skills such as spelling. Also, where a pattern exists, there tends to be more interference for girls than for boys, with the exception

---

[1] The authors wish to express appreciation to Mr. David S. Palermo for his valuable work in helping to secure and process the data on which this paper is based; to Mr. Buford Garner, Superintendent; Mr. David Stewart, Coordinator of Elementary Curriculum; and Mrs. Sylvella Jacobsen, Psychologist, all of the Iowa City Public Schools; and the 15 Iowa City classroom teachers who administered tests, as well as the four elementary school principals concerned, for their help in making this study possible.

CHILD DEVELOPMENT, Vol. 27, No. 3 (September, 1956)

TABLE 1

CORRELATIONS BETWEEN THE CMAS SCORES AND ACADEMIC ACHIEVEMENT MEASURED BY THE IEPT (SUBSECTIONS AND COMPOSITE SCORE)

Grade, Sex	*Vocabulary*	*Reading Comprehension*	*Spelling*	*Arithmetic Computation*	*Composite*
4 B	.00 (53)	—.10 (53)	—.03 (52)	—.13 (52)	—.11 (51)
4 G	—.20 (56)	—.26* (56)	—.23 (57)	—.35** (57)	—.38** (52)
5 B	—.20 (66)	—.08 (66)	—.28* (66)	—.16 (66)	—.18 (66)
5 G	—.11 (60)	—.08 (60)	—.14 (60)	.05 (60)	—.12 (60)
6 B	—.33** (69)	—.33** (69)	—.21 (70)	—.74** (70)	—.32** (68)
6 G	—.58** (51)	—.52** (50)	—.41** (51)	—.57** (52)	—.61** (48)

NOTE.—*N*'s are given in parentheses following each *r*. All *r*'s are Pearsonian.

* Significant at or below the .05 level.

** Significant at or below the .01 level.

of sixth graders and arithmetic. This may be a function of the higher anxiety and L score means for girls or of their generally greater variability (1), and of their stronger academic motivation. It should be added that the CMAS has been useful in predicting performance in complex learning tasks for fourth and fifth graders in other studies (2, 3).

Table 2 shows the same achievement variables' correlations with anxiety scores, but for a population from which has been excluded all children who show L scores of five or higher. (It will be remembered that the L score is high for children who deny that rather universal truths are characteristic of them; who say, for example, "I never lie," or "I always have good manners.")

The correlations in Table 2 show patterns almost identical to those in Table 1. The exclusion of high L-score subjects results in slight reductions in precision of predicting academic success for fourth grade girls, slight additional precision for sixth graders. The over-all pattern, then, is not consistent and the exclusion of the high L-score subjects, of course, results in a substantial reduction in *N*.

INTELLIGENCE

The Otis results were available for the majority of the sixth grade population. These too were given by classroom teachers, and were administered approximately three months after the administration of the CMAS and IEPT, and after the promotion of the sixth grade to junior high school, with some resultant loss of *N*.

For the population of 55 sixth grade boys so tested, the Pearsonian *r* between the Otis and the CMAS was —.16, nonsignificant; for the 45 sixth grade girls, this *r* was —.43, significant at less than the .01 level.

TABLE 2

CORRELATIONS BETWEEN THE CMAS AND ACADEMIC ACHIEVEMENT MEASURED BY THE IEPT (SUBSECTIONS AND COMPOSITE SCORE), WITH HIGH CMAS L SCORES REMOVED FROM THE POPULATION

Grade, Sex	*Vocabulary*	*Reading Comprehension*	*Spelling*	*Arithmetic Computation*	*Composite*
4 B	.00 (33)	.00 (33)	—.01 (33)	.01 (33)	.00 (32)
4 G	—.21 (39)	—.23 (39)	—.19 (39)	—.34* (38)	—.35* (36)
5 B	—.22 (58)	—.10 (58)	—.29* (58)	—.18 (58)	—.22 (58)
5 G	—.06 (48)	—.09 (48)	—.04 (48)	.19 (48)	—.10 (48)
6 B	—.37**(63)	—.36** (63)	—.25* (64)	—.80** (64)	—.34**(62)
6 G	—.58**(47)	—.52** (46)	—.42**(48)	—.58** (48)	—.65**(45)

NOTE.—*N*'s are given in parentheses following each *r*. All *r*'s are Pearsonian.

* Significant at or below the .05 level.

** Significant at or below the .01 level.

Partial *r*'s were computed for the composite score on the IEPT and anxiety, with intelligence held constant. This computation revealed a partial *r* of —.28 for boys, significant at less than the .05 level; and of —.45 for girls, significant at less than the .01 level.

ANXIETY, ACHIEVEMENT AND INTELLIGENCE

The relationships reported above suggested to the authors that the anxiety score might be useful in the practical task of predicting school achievement. Hence, multiple correlations of anxiety and intelligence with the composite score on the IEPT were computed. For this population of sixth graders, *r* between Otis and the composite IEPT for the 55 boys was .56, for the 45 girls, .79. Recomputations of the relationship between anxiety score and composite IEPT score for these populations showed an *r* of —.32 for boys, —.59 for girls (nonsignificant shifts from the larger original population).

The resulting multiple *R*'s are, for boys, .61, and for girls, .84. These represent in each case a small improvement (.05) in prediction from what can be done with Otis score alone.

SUMMARY AND CONCLUSIONS

Both the anxiety and the L score from the children's form of the manifest anxiety scale were found to be related to school achievement, most strongly for the sixth grade portion of a fourth, fifth and sixth grade public school population. The anxiety score was also significantly related to intelligence for sixth grade girls, but, for both sixth grade girls and boys, it retained significant relationships with school achievement when intelligence was partialled out. A small contribution to prediction of academic achievement by the anxiety score, over and above the predictive efficiency of intelligence alone, was found for sixth grade boys and girls.

References

1. Castaneda, A., McCandless, B. R., & Palermo, D. S. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, 27, 317-326.
2. Castaneda, A., Palermo, D. S., & McCandless, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Develpm.*, 1956, 27, 327-332.
3. Palermo, D. S., Castaneda, A., & McCandless, B. R. The relationship of anxiety in children to performance in a complex learning task. *Child Develpm.*, 1956, 27, 333-337.
4. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.

CHILD DEVELOPMENT

Vol. 27
December, 1956
No. 4

Editor
WILLIAM E. MARTIN
Purdue University

Consulting Editors

REGINALD MACGREGOR ARCHIBALD
Rockefeller Institute for Medical Research

JOSEPH H. BRITTON
Pennsylvania State University

SIDNEY W. BIJOU
University of Washington

VAUGHN J. CRANDALL
Fels Research Institute

ERCEL S. EPPRIGHT
Iowa State College

C. GLEN KING
Nutrition Foundation

BENJAMIN PASAMANICK
Ohio State University

ANSELM STRAUSS
University of Chicago

RALPH L. WITHERSPOON
Florida State University

Managing Editor
KATE HOFFMAN

Published Quarterly by

CHILD DEVELOPMENT PUBLICATIONS
(of The Society for Research in Child Development, Inc.)
Purdue University
Lafayette, Indiana

Made in United States of America

Ednl. [illegible] Research
[illegible] COLLEGE
Dated ....
Acc. No. ....

NOTICE TO SUBSCRIBERS

For all new subscriptions and renewals for 1957 and succeeding years received after September 15, 1956, the following rates will apply: CHILD DEVELOPMENT—$8.00 domestic; $8.50 foreign; single issues, $2.50. CHILD DEVELOPMENT ABSTRACTS & BIBLIOGRAPHY—$6.00 domestic; $6.50 foreign; single issues, $2.50. MONOGRAPHS of the SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT—$6.00 domestic; $6.50 foreign; single issue, price varible. CLUB SUBSCRIPTION to all three publications—$16.00 domestic; $17.50 foreign.

Printed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

ANXIETY IN CHILDREN AND SOCIAL STATUS

Boyd R. McCandless,[1] Alfred Castaneda[1]
Iowa Child Welfare Research Station

and David S. Palermo[1]
Southern Illinois University

There have been several studies (e.g., 1, 2, 3, 6, 7, 8, 9, 15, 16, 21) since 1937 dealing with the connection between adjustment and sociometric status in child populations. All have provided at least some support for the hypothesis of a moderate positive relationship between the two ("better adjusted" children are more popular), although Northway and Wigdor (16) find some evidence of a curvilinear relationship. A recent review of sociometric validity (14) summarized a number of studies of adults that also lend support to this hypothesis.

In the child field, the hypothesis has been most recently (and perhaps most rigorously) supported by Thorpe's (19) research with British school children. He studied 980 children in 34 classes, mean chronological age 12-8 years, using a scale that has been shown to possess some validity which he adopted from Eysenck. This test includes seven different "scales" which, grouped, are used to define "neuroticism." Thorpe obtained a pooled *r* of —.152, standard error .034, between "neuroticism" and sociometric status.

In the other child studies mentioned above, the adjustment measure has generally been a global inventory such as the California Test of Personality (7) or the Mental Health Analysis, Elementary or Intermediate Series, Form A (1, 6). Rorschachs have also been used (16), as well as

---

[1] The authors wish to express appreciation to Mr. Buford Garner, Superintendent; Mr. David Stewart, Coordinator of Elementary Curriculum; and Mrs. Sylvella Jacobsen, Psychologist, all of the Iowa City Public Schools; and the 15 Iowa City classroom teachers who administered tests, as well as the four elementary school principals concerned, for their help in making this study possible.

CHILD DEVELOPMENT, Vol. 27, No. 4 (December, 1956)

teachers' or observers' ratings, interviews, case histories (2, 3, 8, 15) and problem check lists (9).

Previous studies by the present authors have reported the form, scoring, norms and reliabilities (4) for the children's form of the manifest anxiety scale (CMAS), as well as data concerning relationships of the scale with performance in complex learning situations (5, 17) and in the area of academic achievement (12). The literature demonstrating predictive power for Taylor's adult form of the manifest anxiety scale (18) is voluminous. The authors have also been interested for some time in sociometric techniques from the points of view both of their reliability and usefulness (10, 11).

Table 1

TEST-RETEST RELIABILITIES (ONE-WEEK INTERVAL, PRODUCT-MOMENT) FOR RATE AND RANK METHODS OF ASSESSING SOCIOMETRIC STATUS, BY GRADE GROUPS AND BY SEX

	CLASS									
	1		2		3		4		5	
Grade, Sex	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*	*N*	*r*
4 B	8	(r).93**	15	(r).75**	9	.77*	10	.92**	15	.99**
4 G	11	(r).98**	10	(r).70*	15	.90**	15	.92**	10	.97**
5 B	10	.93**	17	.95**	11	.95**	16	(r).96**	16	(r).55*
5 G	12	.80**	16	.96**	13	.96**	7	(r).91**	16	(r).96**
6 B	11	.95**	18	.90**	13	(r).94**	7	(r).91**	18	.94**
6 G	5	.91*	12	.98**	10	(r).94**	11	(r).82**	12	.96**

NOTE.—An (r) preceding a coefficient indicates that it was obtained by the rate method.

* Significant at less than the .05 level.

** Significant at less than the .01 level.

SOCIOMETRIC PROCEDURES AND RESULTS

Two teacher-administered sociometric techniques were administered, the rank method to nine classes, the rate method to six classes of public school fourth, fifth and sixth grade children. These were repeated, same methods for same classes, one week later to determine test-retest reliability. The population taking both tests numbered 369 children, 194 boys and 175 girls. The sociometric was a one-question "friend" instrument. This choice of a one-question, rank and rate sociometric was due to the authors' quest for a very simple instrument which could be administered by teachers from prepared instruction and data sheets.

The rank method was administered (by sexes rather than whole classes) in this fashion: a mimeographed list of all girls in a class was distributed to each girl, a list of all boys to each boy in a class. Children were instructed to enter a (1) after the name of their very best friend, a (2) after the name of their second best friend, and so on until every child but the one doing the ranking had a different number after his name, the largest number (*n*) being the number of boys or girls in a class, excepting the child doing the ranking.

The rate method was also given by sex. A sheet similar to the rank sheet was handed to each boy or girl, with the names of all the boys or girls in a class in a column down the left hand side of the page. A line to the right of each name included five equidistant check points, each subsumed by a number from (1) to (5) reading from left to right. The child was instructed to check a point for each name other than his own, a (1) if the child was a "best friend," a (3) if he didn't know the child very well or neither liked nor disliked him, and a (5) if the child "is not my friend," etc.

Table 1 gives reliability figures (product moment r's) by class, grade and sex for the two methods of sociometric assessment. The scores on which these correlations are based are the average rating or the average ranking received by a child. This score was the average of all ranks or ratings given by other children of his own sex to a given child. Consequently the lowest score represents the most popular child for a sex-class group. However, for clearness of presentation, signs of correlation coefficients in Table 2 above have been changed so that a negative correlation means that there is a tendency for the more anxious children to be less popular, and vice versa.

Table 2

POOLED PRODUCT-MOMENT CORRELATIONS BY GRADE AND SEX BETWEEN CMAS SCORES AND SOCIOMETRIC STATUS

Grade, Sex	*N Classes*	*N Individuals*	*r*
4 B	5	58	—.28*
4 G	5	62	—.23
5 B	5	72	—.51**
5 G	5	69	—.75**,†
6 B	5	73	—.16
6 G	5	53	.01

* Significant at less than the .05 level.

** Significant at less than the .01 level.

† This is an estimate based on five fifth grade classes, the correlation population of which proved to be nonhomogeneous at just below the .05 level (all other class populations were homogeneous). The correlations and N's for the five classes were: 1. $r = —.90$, $N = 14$; 2. $r = —.55$, $N = 17$; 3. $r = —.46$, $N = 15$; 4. $r = +.03$, $N = 7$; 5. $r = —.81$, $N = 16$.

There is no clearcut reliability difference between the rate and the rank methods; of the four reliabilities below .80, three were obtained by use of the rate method, two of these three being for fourth grade groups; and of the 12 reliabilities of .95 or more, only three were obtained by the rate method. Only four of the 30 reliabilities computed fail to reach significance at less than the .01 level. Two of these were obtained by the rank, two by the rate method, and all four were significant at less than the .05 level. Median test-retest reliability for the rate method (12 measures for six classes) was .92; and for the rank method (18 measures for nine classes) was .95. However, correlation range for the rate method (.55–.98) was greater than for the rank method (.77–.99).

At a more general level, this study indicates that a single question, teacher-administered-by-group sociometric, sexes separated, results in high reliabilities as judged by test-retest at a one-week interval. The reliabilities compare favorably with the studies reviewed by either Mouton, Blake and Fruchter (13) or by Witryol and Thompson (20).

ANXIETY SCALE PROCEDURES AND RESULTS

The hypothesis for this section of the study was that there would be a negative relationship between anxiety and social acceptability in fourth, fifth and sixth grade public school children (i.e., that more anxious children would be less popular). Subjects were 387 fourth, fifth and sixth graders from the public schools of a midwestern town of about 27,000 population. There were 203 boys and 184 girls. The population completely overlaps the sociometric population described above, differences in *N* being due to absences on retests for the sociometric. All relationships reported here are between a first administration of the CMAS and the sociometric assessment. Rank and rate methods of collecting sociometric data are combined in this portion of the study, since there were no significant differences in their relationship with CMAS scores between the two methods.

Table 2 shows correlations by grade and sex of the CMAS scores and sociometric status. The CMAS, as was also true of the sociometric assessments, was administered by classroom teachers. It was also repeated a week later to provide the reliability data reported in (4).

The over-all picture from Table 2 supports the hypothesis of a negative relationship between anxiety and social status (more anxious children are less popular) although there is clear variation by grades. The reasonably substantial *N*'s, both of classes and individuals, the relative consistency by sexes in a grade (see below) and the frequency of significant correlations all argue against the notion of random variation as an explanation of this correlation pattern. Fourth graders, boys and girls, have low negative *r*'s (for fourth grade boys *r* is significant at less than the .05 level, and *r* for fourth grade girls misses the .05 significance level by only .02). Correlations for both fifth grade boys and girls are significant at less than the .001 level, while sixth graders' *r*'s hover close to zero. A social class differential cannot be argued, since the fourth, fifth and sixth graders are drawn by groups, each group coming from the same school. One can only suspect different social interaction patterns which are a function of unknown factors, possibly of grade and age (including physical maturity), teacher influence, or other, and look further for reasons for the variability.

As a check on within-class similarity (for which teachers might possibly be thought to be most responsible) an analysis of variance was run to test the relationship of within-groups variance (within a given class) to between groups variance (all grades and classes). Utilizing the 15 classes, including the one fifth grade class of girls (class 1, footnote, Table 2) that contributed most heavily to nonhomogeneity, the resulting F was 1.80, which is nonsignificant for 14 and 15 *d.f.* ($F = 2.43$ is required for .05 level of significance). Eliminating this class, F equalled 3.60, which is significant for 13 and 14 *d.f.* at between the .01 and .02 levels. This analysis suggests very tentatively that there is a tendency for a given class, regardless of sex, to follow a given pattern of anxiety-social status relationship. The over-all nonhomogeneity of the correlations in Table 2 lends further support to this idea.

However, this over-all nonhomogeneity (the null hypothesis of no difference in the total correlation population is rejected at less than the .01 level) is due most heavily to the single class of fifth grade girls mentioned above. Dropping this single half-class enables us to retain the null hypothesis at between the .20 and .10 levels, and results in an average correlation between anxiety and social acceptance of —.32, significant at less than the .001 level. Average *r*'s for boys and girls respectively (dropping the single deviant class of fifth grade girls) are —.33 and —.30, both significant at less than the .01 level.

Summary and Conclusions

High reliabilities, using the test-retest method at a one-week interval, were obtained for single-question friendship sociometrics administered by classroom teachers. The rank and rate methods were administered to fourth, fifth and sixth grade populations, the sexes separated, to obtain these reliabilities. There is a suggestion that the rank method is slightly more reliable, although differences in reliabilities between the two methods are slight.

Relationships between scores obtained by using the children's form of the manifest anxiety scale, and sociometric status, were predominantly negative (i.e., the more anxious youngsters were the less popular) for both boys and girls. These relationships were high for fifth graders, moderate for fourth graders and approximately zero for sixth graders. There is some indication that there may be characteristic "class climates" not dependent upon differences between the sexes. This, at a speculative level, may possibly be said to be related to the influence of the teacher on the class. An over-all lack of homogeneity of the correlation population supports the notion of "class climate," although when one half class of deviant fifth grade girls is dropped, the correlation population shows homogeneity. The resulting average correlation between anxiety and sociometric standing is then —.32, statistically significant. The authors believe, however, that looking at the correlations by grade and sex provides a more meaningful way of regarding the data than does concentration on the average *r*.

REFERENCES

1. BARON, DENIS. Personal-social characteristics and classroom social status. *Sociometry,* 1951, 14, 32-42.
2. BONNEY, MERL E. A study of social status on the second grade level. *J. genet. Psychol.,* 1942, 60, 271-305.
3. BONNEY, MERL E. Popular and unpopular children, a sociometric study. *Sociometry Monogr.,* 1947, No. 9.
4. CASTANEDA, ALFRED, MCCANDLESS, BOYD R., & PALERMO, DAVID S. The children's form of the manifest anxiety scale. *Child Develpm.,* 1956, 27, 317-326.
5. CASTANEDA, ALFRED, PALERMO, DAVID S., & MCCANDLESS, BOYD R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Develpm.,* 1956, 27, 327-332.
6. GREENBLATT, E. L. Relationship of mental health and social status. *J. educ. Res.,* 1950, 44, 193-204.
7. GROSSMAN, BEVERLY, & WRIGHTER, JOYCE. The relationship between selection-rejection and intelligence, social status, and personality among sixth grade children. *Sociometry,* 1948, 11, 346-355.
8. HARDY, MARTHA C. Social recognition at the elementary school age. *J. soc. Psychol.,* 1937, 8, 365-384.
9. KUHLEN, RAYMOND G., & BRETSCH, HOWARD S. Sociometric status and personal problems of adolescents. *Sociometry,* 1947, 10, 122-132.
10. MCCANDLESS, BOYD R. Changing relationships between dominance and social acceptability during group democratization. *Amer. J. Orthopsychiat.,* 1942, 12, 529-536.
11. MCCANDLESS, BOYD R. Motivational and social structure research at pre- and elementary school levels. Presidential address before

Division 7, American Psychological Association, San Francisco, September, 1955.

12. McCandless, Boyd R., & Castaneda, Alfred. Anxiety in children, school achievement and intelligence. *Child Develpm.*, 1956, 27, 379-382.

13. Mouton, Jane S., Blake, Robert R., & Fruchter, Benjamin. The reliability of sociometric measures. *Sociometry*, 1955, 18, 7-48.

14. Mouton, Jane S., Blake, Robert R., & Fruchter, Benjamin. The validity of sociometric responses. *Sociometry*, 1955, 18, 181-206.

15. Northway, Mary L. Outsiders; a study of the personality patterns of children least acceptable to their age mates. *Sociometry*, 1944, 7, 10-25.

16. Northway, Mary L., & Wigdor, Blossom T. Rorschach patterns related to sociometric status of school children. *Sociometry*, 1947, 10, 186-199.

17. Palermo, David S., Castaneda, Alfred, & McCandless, Boyd R. The relationship of anxiety in children to performance in a complex learning task. *Child Develpm.*, 1956, 27, 333-337.

18. Taylor, Janet A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.

19. Thorpe, J. G. An investigation into some correlates of sociometric status within school classes. *Sociometry*, 1955, 18, 49-61.

20. Witryol, Sam L., & Thompson, George G. A critical review of the stability and social acceptability scores obtained with the partial-rank-order and the paired comparison scales. *Genet. Psychol. Monogr.*, 1953, 48, 221-260.

21. Young, Lyle L., & Cooper, Dan H. Some factors associated with popularity. *J. educ. Psychol.*, 1944, 35, 513-535.

SOME EMOTIONAL ATTITUDES OF THE YOUNG CHILD IN RELATION TO CHARACTERISTICS OF HIS SIBLING[1,2]

HELEN L. KOCH
University of Chicago

The study is part of a more extensive one dealing with the question of the relation between personality characteristics of the five- and six-year-old from two-child families and his ordinal position, the sex of his sibling, and the age difference which separates him from his sibling. The dependent variables to be related here to the independent variables named above are teachers' ratings of the children on the following traits: excitability, intensity of emotional response, speed of recovery from emotional disturbance, stability of mood, nervous habits, physical activeness, health, apprehensiveness relative to physical activities, social apprehensiveness, sensitiveness, tendency to anger readily, self-confidence, finality of decision or degree of vacillation, cheerfulness, tendency to alibi, tendency to project blame, and indirectness of response to fear and frustration.

The subjects were 384 five- and six-year-olds from two-child, urban, white, native-born, intact families. Both sibs were "normal" physically and mentally; i.e., a child with any serious defect or with a sib who so suffered

---

[1] The investigation was supported, in part, by a grant from the National Institute of Mental Health, of the National Institutes of Health, U.S. Public Health Service, and, in part, by the Sonia Shankman Orthogenic School Research Fund, University of Chicago.

[2] The author wishes to acknowledge the help received from, and to express her gratitude to, the teachers and principals of the following schools in Chicago which participated in the study: Bennett School, Bradwell School, Field School, Holmes School, Gunsalus School, Luella School, Perry School, Schmid School, Sullivan School, Wentworth School, University of Chicago Laboratory Schools, and the University of Chicago Nursery School. The author is also much indebted to Dr. Grace Munson and Dr. Thelma Thurstone, both formerly Directors of the Bureau of Child Study, Chicago Public Schools. Dr. Helen Heath did most of the statistical work for the study. We acknowledge with gratitude her discriminating and loyal service.

CHILD DEVELOPMENT, Vol. 27, No. 4 (December, 1956)

was eliminated from the study. No foster children were included and, as far as the school knew, the family was intact. If, in the interview we had with the child, we discovered we had reason to suspect the family members might not be living together, we dropped the child as a subject.

The children were grouped into 24 matched subgroups of 16 each. The design involved child subjects of two sexes, siblings of two sexes, two ordinal positions, and three age-difference levels. There were 48 children in each of the following groups: male with a male sib older, male with a male sib younger, male with a female sib older, male with a female sib younger, female with a male sib older, female with a male sib younger, female with a female sib older, female with a female sib younger. Each of the groups of 48 were divided into three subgroups of 16 each; the children in the three subgroups differing in age from their sib by (1) less than two years, (2) two to four years, and (3) four to six years, respectively. The children were drawn chiefly from the public schools of Chicago, only one private school system, the University of Chicago Laboratory Schools, contributing subjects.

The matching was done individual by individual on the basis of age, and the socio-economic status of the child's father's occupation (5), as well as the neighborhood of residence of the child's family (13). Matching, of course, could not be exact but 98 per cent of the matchings in age were within six months; 93 per cent, within one level on the occupational scale; and 88 per cent, within one level on the neighborhood scale. Level 4 on the Minnesota occupation scale (5) was omitted because the children were all urban residents. The group compositions have been described in detail in an earlier publication in this journal (9).

We have no direct reliability measure of the ratings but, although our measures are not the same and the populations very different, the general level of the reliabilities may perhaps be estimated *roughly* from those presented in References (2) and (14). There is, of course, a question of the extent to which our measures are fogged by halo effects and the teachers' inability to discriminate differences between children and between traits. The teachers had worked with most of the children they rated from five to nine months, three months being the lower limit of duration of contact permitted. The ratings were, of course, based chiefly on school behavior and were apparently conscientiously made. Since relatively more of our subjects in the more difficult-to-locate family constellation patterns came from two very large schools, it is possible a school standard factor may be blurring results some. We suspect this effect is not considerable, as the teachers were given in each rating scale a rather careful definition of each trait, and most of the teachers had taught in a number of schools. Hence their standards of child behavior were not based solely on their contact with the children of one school. The matching of groups on a socio-economic basis probably further diluted the effect of a school-standard variable, if one were operative. The individual judge variable was probably fairly well controlled, too, because there were many raters and their ratings were well scattered through the groups. Most teachers rated only a few cases. There are other possible selective effects in our data. For a discussion of some of these see References (8) and (9).

The rating scales used were drawn from the Fels Child Behavior Rating Scales (14) and the California Behavior Inventory for Nursery School Children (1). The Fels scales are indicated by a star next to the caption in the tables of means, the unstarred scales, then, being the California ones. All scales were treated as line scales. The teachers' line checkings were converted into ratings on a nine-point scale with the help of a grid and these ratings were then normalized on the basis of a population of 498 five- and six-year-olds, of which our groups constituted the major part.

The Bartlett test (3, pp. 195-200) was applied to determine whether there was any evidence the subgroups might be drawn from populations with different variances. There was no evidence they were.

The analysis of variance data for the total population only are offered in Table 10; and, if a trait showed no significant relation to any of our independent variables, no variance data at all are presented. The traits of which the latter was true are: intensity of emotional response, physical apprehensiveness, moodiness, emotional excitability, and tendency to project blame. *F*'s are given in Table 10 only if the group difference is significant at the 5 per cent point or better. To determine significance of the differences between various subgroups, *t*'s were computed when indicated but are not offered here in the interest of economy. Unless specified, the differences labelled significant in the text are so at the 5 per cent point at least.

The reader is referred to References (1) and (14) for the definitions of the various traits studied. In the main, the definitions are traditional and need not be reproduced here. The "health" scale we are somewhat uneasy about because our teachers probably were not well qualified to make a judgment on this trait. Their impression of a child's health was doubtless influenced by the latter's attendance record (which, of course, may be as much a function of the parents' attitudes as of the child's health), the parents' comments on the child, as well as the appearance and behavior of the child himself. The health ratings, hence, will be difficult to interpret. Moreover, whether we can assume to be constant from group to group any error there is, is also a question. We suspect, for instance, parents' attitudes toward health and school attendance may be a function in part of the variables studied. It seems not too unlikely that parents would be more protective of their first-born than later children or more rigidly insistent on regular attendance for him; or that one sex would be kept out of school more than the other. We thought for a while we would employ school attendance records in the place of health ratings, but the former were found completely unusable, being subject to all the hazards of the ratings employed, as well as those stemming, among other things, from arbitrary school regulations relative to when the child was or was not on the school rolls.

The definition of sensitiveness in the Fels Scale we suspect would cause the rating on this trait to be strongly influenced by an alertness characteristic in the child as well as by signs that the child's feelings were easily hurt. This two-dimensional characteristic of the defined trait may make the interpretations of the ratings difficult. A fairly confident child may be very alert to the relevance of situations for his status, for instance, without being particularly easily hurt. How the two vectors combine in the ratings is not very clear. There is evidence some of the Fels judges were unable to agree well on this trait (14).

The nervous habits listed in our definition of these include such items as thumb-sucking, nail biting, blinking, and grimacing. The magnitude of the rating is a function of both the number of different types of habits exhibited as well as of the frequency of occurrence of any type. Just what these behavior patterns mean in the individual's adjustments is not too clear. The list does not include any visceral activities which actually may be more diagnostic of tension than the items mentioned. The overt patterns may be highly correlated with energy level or with activeness denied more constructive outlets in the classroom. Local irritants instigating the automatisms may also vary from group to group, being, for instance, possibly greater for boys than for girls who are less active and are bruised less. Schoolroom atmosphere may be another variable influencing the number and frequency of occurrence of the nervous habits shown, but, since many classes contributed to the populations of our various subgroups, classroom atmosphere is probably no important determinant of our group differences. Let us add the comment that Conrad's judges did not agree well in their ratings of the trait, nervous habits (2).

Since the relations to be described are complicated, we shall use the outline form in characterizing them, first describing relations in terms of our various independent variables. Later we shall offer a summary characterization of the various sib's-sex and ordinal-position groups. Also, since many interactions are significant, the same relation will be listed under two or three headings, though, in each, its description may be given a different emphasis.

To save space we shall often refer to the groups in code. The first letter and number in the code will indicate the sex and ordinal position of the children in the group; the second letter, the sex of their siblings; and the numbers in parentheses, the range of the difference in age (in months) between the child subjects and their sibs. Thus M1F(25-48) refers to the group of first-born boys each of whom has one sib, a sister younger by 25 to 48 months.

TABLE 1

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: EMOTIONAL EXCITABILITY* AND INTENSITY OF EMOTIONAL RESPONSE*

Subgroup			Emotional Excitability*				Intensity of Emotional Response*			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	4.209	4.181	3.859	4.083	3.922	3.938	4.170	4.010
M	2	F	4.188	3.642	4.211	4.014	3.867	3.738	4.111	3.905
F	2	M	4.086	4.102	3.919	4.036	3.900	3.972	3.951	3.941
F	2	F	4.044	3.982	4.157	4.061	4.012	3.886	4.164	4.021
M	1	M	3.767	4.100	3.769	3.879	3.734	4.409	4.191	4.111
M	1	F	3.898	4.469	4.043	4.137	3.559	4.363	3.963	3.962
F	1	M	4.005	3.960	3.985	3.983	4.285	4.023	3.917	4.075
F	1	F	3.792	4.099	4.216	4.036	3.788	4.187	4.434	4.137
M	2	M or F	4.199	3.912	4.035	4.049	3.894	3.838	4.140	3.958
F	2	M or F	4.065	4.042	4.038	4.048	3.956	3.929	4.058	3.981
M	1	M or F	3.832	4.285	3.906	4.008	3.646	4.386	4.077	4.036
F	1	M or F	3.898	4.030	4.101	4.010	4.037	4.105	4.176	4.106
M or F	2	M	4.148	4.142	3.889	4.059	3.911	3.955	4.061	3.975
M or F	2	F	4.116	3.813	4.184	4.038	3.939	3.812	4.138	3.963
M or F	1	M	3.886	4.030	3.877	3.931	4.009	4.216	4.054	4.093
M or F	1	F	3.845	4.284	4.129	4.086	3.673	4.275	4.198	4.049
M	1 or 2	M	3.988	4.141	3.814	3.981	3.828	4.173	4.181	4.061
M	1 or 2	F	4.043	4.056	4.127	4.075	3.713	4.051	4.037	3.933
F	1 or 2	M	4.045	4.031	3.952	4.010	4.093	3.998	3.934	3.008
F	1 or 2	F	3.918	4.041	4.187	4.048	3.900	4.036	4.299	4.079
M	1 or 2	M or F	4.016	4.098	3.971	4.028	3.770	4.112	4.109	3.997
F	1 or 2	M or F	3.982	4.036	4.069	4.029	3.996	4.017	4.117	4.043
M or F	1 or 2	M	4.017	4.086	3.883	3.995	3.960	4.085	4.057	4.034
M or F	1 or 2	F	3.980	4.048	4.157	4.062	3.806	4.044	4.168	4.006
M or F	2	M or F	4.132	3.977	4.037	4.048	3.925	3.883	4.099	3.969
M or F	1	M or F	3.865	4.157	4.003	4.009	3.841	4.246	4.126	4.071
M or F	1 or 2	M or F	3.999	4.067	4.020	4.029	3.883	4.064	4.113	4.020

Ordinal-Position-Group Differences

1. Second-borns were rated significantly higher in speed and degree of recovery from emotional upsets than were first-borns only when the sib spacing was under two years (Tables 2, 10).

2. The younger sib among males was judged to have more nervous habits or to indulge in them more frequently than the older sib, while the opposite trend relative to ordinal position obtained for females (Tables 8, 10).

3. The health estimates for first-borns from opposite-sex sib pairs tended to be better than those for second-borns from parallel groups;

TABLE 2

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: SPEED OF RECOVERY FROM EMOTIONAL DISTURBANCE AND MOODINESS

Subgroup			Speed of Recovery from Emotional Disturbance				Moodiness			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	3.940	3.958	4.311	4.070	4.081	4.089	3.958	4.043
M	2	F	4.211	4.298	3.824	4.111	4.159	3.911	4.124	4.065
F	2	M	4.468	4.130	3.695	4.098	4.111	3.848	3.828	3.929
F	2	F	4.023	3.868	3.966	3.952	3.896	4.079	3.553	3.843
M	1	M	4.188	3.495	4.009	3.897	3.993	4.552	3.768	4.104
M	1	F	3.653	4.016	4.099	3.923	4.116	4.386	4.175	4.226
F	1	M	3.318	4.061	4.201	3.860	4.108	3.839	4.271	4.073
F	1	F	3.880	3.774	4.191	3.948	4.417	4.104	3.869	4.130
M	2	M or F	4.076	4.128	4.068	4.090	4.120	4.000	4.041	4.054
F	2	M or F	4.245	3.999	3.830	4.025	4.004	3.963	3.690	3.886
M	1	M or F	3.920	3.755	4.054	3.910	4.054	4.469	3.972	4.165
F	1	M or F	3.599	3.917	4.196	3.904	4.263	3.971	4.070	4.101
M or F	2	M	4.204	4.044	4.003	4.084	4.096	3.968	3.893	3.986
M or F	2	F	4.117	4.083	3.895	4.032	4.028	3.995	3.839	3.954
M or F	1	M	3.753	3.778	4.105	3.879	4.050	4.195	4.020	4.088
M or F	1	F	3.767	3.895	4.145	3.935	4.267	4.245	4.022	4.178
M	1 or 2	M	4.064	3.727	4.160	3.984	4.037	4.320	3.863	4.073
M	1 or 2	F	3.932	4.157	3.961	4.017	4.138	4.149	4.150	4.145
F	1 or 2	M	3.893	4.095	3.948	3.979	4.110	3.843	4.049	4.001
F	1 or 2	F	3.952	3.821	4.078	3.950	4.157	4.091	3.711	3.986
M	1 or 2	M or F	3.998	3.942	4.061	4.000	4.087	4.234	4.006	4.109
F	1 or 2	M or F	3.922	3.958	4.013	3.965	4.133	3.967	3.880	3.994
M or F	1 or 2	M	3.978	3.911	4.054	3.981	4.073	4.082	3.956	4.037
M or F	1 or 2	F	3.942	3.989	4.020	3.984	4.147	4.120	3.930	4.066
M or F	2	M or F	4.161	4.063	3.949	4.058	4.062	3.981	3.866	3.970
M or F	1	M or F	3.760	3.836	4.125	3.907	4.158	4.220	4.021	4.133
M or F	1 or 2	M or F	3.960	3.950	4.037	3.982	4.110	4.101	3.943	4.052

while, in the case of those children from same-sex sib pairs, the ordinal-position trend was reversed (Tables 9, 10).

4. Second-born males at the closest spacing earned a higher mean rating in the trait, tendency to anger, than did first-born males; but at the middle spacing first-born males rated higher than did second-born. The trend among females was for first-borns to exceed second-borns, the difference being significant, however, only at the widest spacing (Tables 4, 10).

5. Those children who were the older in their sib pairs tended to receive a higher rating in self-confidence than did those who were the younger member (Tables 5, 10).

TABLE 3

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: SOCIAL APPREHENSIVENESS* AND PHYSICAL APPREHENSIVENESS*

Subgroup			Social Apprehensiveness*				Physical Apprehensiveness*			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	4.424	4.154	3.697	4.092	3.946	3.924	3.766	3.879
M	2	F	4.346	3.974	4.091	4.137	3.832	4.006	4.074	3.971
F	2	M	4.091	3.802	4.067	3.987	3.928	3.671	4.098	3.899
F	2	F	3.849	4.306	4.014	4.056	3.558	4.079	4.040	3.892
M	1	M	4.148	4.143	3.765	4.019	4.485	3.759	4.317	4.187
M	1	F	4.328	3.812	4.066	4.069	3.971	3.677	4.285	3.978
F	1	M	3.814	3.793	3.450	3.685	3.823	4.047	3.901	3.923
F	1	F	3.983	4.103	4.013	4.033	4.175	3.929	3.988	4.031
M	2	M or F	4.385	4.064	3.894	4.115	3.889	3.965	3.920	3.925
F	2	M or F	3.970	4.054	4.040	4.021	3.743	3.875	4.069	3.896
M	1	M or F	4.238	3.977	3.916	4.044	4.228	3.718	4.301	4.082
F	1	M or F	3.898	3.948	3.732	3.859	3.999	3.988	3.944	3.977
M or F	2	M	4.258	3.978	3.882	4.039	3.937	3.798	3.932	3.889
M or F	2	F	4.098	4.140	4.053	4.097	3.695	4.043	4.057	3.932
M or F	1	M	3.981	3.968	3.608	3.852	4.154	3.903	4.109	4.055
M or F	1	F	4.155	3.957	4.040	4.051	4.073	3.803	4.137	4.004
M	1 or 2	M	4.286	4.148	3.731	4.055	4.215	3.842	4.042	4.033
M	1 or 2	F	4.337	3.893	4.079	4.103	3.902	3.841	4.180	3.974
F	1 or 2	M	3.952	3.797	3.758	3.836	3.875	3.859	3.999	3.911
F	1 or 2	F	3.916	4.204	4.013	4.045	3.866	4.004	4.014	3.962
M	1 or 2	M or F	4.312	4.021	3.905	4.079	4.058	3.841	4.111	4.004
F	1 or 2	M or F	3.934	4.001	3.886	3.940	3.871	3.932	4.007	3.936
M or F	1 or 2	M	4.119	3.973	3.745	3.946	4.045	3.850	4.020	3.972
M or F	1 or 2	F	4.127	4.049	4.046	4.074	3.884	3.923	4.097	3.968
M or F	2	M or F	4.178	4.0[illegible]9	3.967	4.069	3.816	3.920	3.995	3.910
M or F	1	M or F	4.068	3.962	3.824	3.952	4.113	3.853	4.123	4.030
M or F	1 or 2	M or F	4.1[illegible]	4.011	3.895	4.010	3.965	3.887	4.059	3.970

6. Girls with a younger sister were judged more vacillating than girls with an older sister when the children differed from their sibs by over four years in age or by less than two (Tables 5, 10).

7. First-borns at the two-to-four-year sib-age-difference level were judged to alibi more than second-borns (Tables 6, 10).

8. First-born boys at the middle spacing received a higher mean rating in tendency to respond indirectly to fear and frustration than did second-born; whereas, at the same spacing, girls with an older sister scored higher than did girls with a younger (Tables 7, 10).

TABLE 4

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: CHEERFULNESS* AND TENDENCY TO ANGER

Subgroup			Cheerfulness*				Tendency to Anger			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	3.609	3.971	4.090	3.890	4.582	3.959	4.291	4.277
M	2	F	3.898	4.198	4.292	4.129	4.191	3.938	4.327	4.152
F	2	M	4.388	4.133	3.913	4.144	4.015	3.704	3.694	3.804
F	2	F	4.184	3.784	4.113	4.027	3.765	3.849	3.763	3.792
M	1	M	3.970	3.426	4.072	3.823	3.758	4.413	4.022	4.064
M	1	F	3.978	3.922	4.006	3.969	3.836	4.603	4.241	4.227
F	1	M	3.871	4.095	3.984	3.984	4.263	3.899	4.239	4.134
F	1	F	3.981	3.550	3.839	3.790	3.767	4.037	4.198	4.000
M	2	M or F	3.753	4.084	4.191	4.010	4.386	3.948	4.309	4.215
F	2	M or F	4.286	3.958	4.013	4.086	3.890	3.777	3.728	3.798
M	1	M or F	3.974	3.674	4.039	3.896	3.797	4.508	4.132	4.145
F	1	M or F	3.926	3.823	3.912	3.887	4.015	3.968	4.218	4.067
M or F	2	M	3.999	4.052	4.001	4.017	4.298	3.831	3.993	4.041
M or F	2	F	4.041	3.991	4.202	4.078	3.978	3.894	4.045	3.972
M or F	1	M	3.921	3.761	4.028	3.903	4.010	4.156	4.131	4.099
M or F	1	F	3.980	3.736	3.923	3.880	3.802	4.320	4.219	4.114
M	1 or 2	M	3.790	3.698	4.081	3.856	4.170	4.186	4.156	4.171
M	1 or 2	F	3.938	4.060	4.149	4.049	4.013	4.270	4.284	4.189
F	1 or 2	M	4.130	4.114	3.948	4.064	4.139	3.801	3.967	3.969
F	1 or 2	F	4.083	3.667	3.976	3.909	3.766	3.943	3.980	3.896
M	1 or 2	M or F	3.864	3.879	4.115	3.953	4.092	4.228	4.220	4.180
F	1 or 2	M or F	4.106	3.891	3.962	3.986	3.952	3.872	3.973	3.933
M or F	1 or 2	M	3.960	3.906	4.015	3.960	4.154	3.994	4.062	4.070
M or F	1 or 2	F	4.010	3.864	4.063	3.979	3.890	4.107	4.132	4.043
M or F	2	M or F	4.020	4.021	4.102	4.048	4.138	3.863	4.019	4.006
M or F	1	M or F	3.950	3.748	3.976	3.891	3.906	4.238	4.175	4.106
M or F	1 or 2	M or F	3.985	3.885	4.039	3.970	4.022	4.050	4.097	4.056

Sib's-Sex-Group Differences

9. Children whose sib was opposite in sex, if the former were first-borns and very near the sib in age, were judged to recover more slowly and less adequately from emotional disturbances than children from parallel groups where the sib was the same in sex. At the middle spacing those from opposite-sex sib pairs were rated the higher (Tables 2, 10).

10. When the child and his sib were four to six years apart in age, those subjects from mixed-sex sib pairs were gauged to have the more nervous habits or show them more frequently (Tables 8, 10).

11. The health assessment for first-borns whose sib was not of their sex and was less than four years younger was better than that for the

TABLE 5

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: SELF-CONFIDENCE AND FINALITY OF DECISION

Subgroup			Self-Confidence				Finality of Decision			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	3.806	4.006	4.284	4.032	3.869	3.987	4.463	4.106
M	2	F	3.543	4.033	3.757	3.778	4.103	4.089	3.756	3.983
F	2	M	3.984	4.304	3.856	4.048	4.291	4.721	3.850	4.287
F	2	F	4.119	3.584	4.007	3.904	4.404	3.593	4.251	4.083
M	1	M	3.782	3.854	4.167	3.934	3.855	4.079	4.093	4.009
M	1	F	3.976	4.528	3.996	4.167	4.013	3.947	4.190	4.050
F	1	M	4.432	4.262	4.351	4.348	3.908	4.435	3.886	4.077
F	1	F	4.189	4.078	4.126	4.131	3.846	3.735	3.642	3.741
M	2	M or F	3.675	4.019	4.021	3.905	3.986	4.038	4.109	4.044
F	2	M or F	4.052	3.944	3.932	3.976	4.348	4.157	4.050	4.185
M	1	M or F	3.879	4.191	4.081	4.050	3.934	4.013	4.141	4.029
F	1	M or F	4.311	4.170	4.238	4.240	3.877	4.085	3.764	3.909
M or F	2	M	3.895	4.155	4.070	4.040	4.080	4.354	4.156	4.197
M or F	2	F	3.831	3.808	3.882	3.841	4.253	3.841	4.003	4.033
M or F	1	M	4.107	4.058	4.259	4.141	3.882	4.257	3.989	4.043
M or F	1	F	4.083	4.303	4.061	4.149	3.929	3.841	3.916	3.895
M	1 or 2	M	3.794	3.930	4.226	3.983	3.862	4.033	4.278	4.057
M	1 or 2	F	3.759	4.280	3.876	3.972	4.058	4.018	3.973	4.016
F	1 or 2	M	4.208	4.283	4.103	4.198	4.099	4.578	3.868	4.182
F	1 or 2	F	4.154	3.831	4.066	4.017	4.125	3.664	3.946	3.912
M	1 or 2	M or F	3.777	4.105	4.051	3.978	3.960	4.026	4.125	4.037
F	1 or 2	M or F	4.181	4.057	4.085	4.108	4.112	4.121	3.907	4.047
M or F	1 or 2	M	4.001	4.106	4.165	4.091	3.981	4.306	4.073	4.120
M or F	1 or 2	F	3.957	4.056	3.971	3.995	4.091	3.841	3.960	3.964
M or F	2	M or F	3.863	3.982	3.976	3.940	4.167	4.098	4.080	4.115
M or F	1	M or F	4.095	4.181	4.160	4.145	3.906	4.049	3.953	3.969
M or F	1 or 2	M or F	3.979	4.081	4.068	4.043	4.036	4.073	4.016	4.042

children in parallel groups who had a sib of the same sex; whereas, among second-borns at the closest spacing, those with a sib identical in sex were judged healthier than those with a sib differing in sex (Tables 9, 10).

12. When their sib was male and differed in age by two to four years, boys were judged more apprehensive in social situations than girls; but, when the sib was female, girls were judged the more apprehensive (Tables 3, 10).

13. Children with a male sib were rated higher in sensitiveness than those with a female sib (Tables 7, 10).

14. The subjects whose sib was opposite in sex and either older or younger by two to four years received a higher mean rating in self-confi-

TABLE 6

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: TENDENCY TO PROJECT BLAME AND TENDENCY TO ALIBI

Subgroup			Tendency to Project Blame				Tendency to Alibi			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7 - 24	25 - 48	49 - 72	Total	7 - 24	25 - 48	49 - 72	Total
			Score Means				Score Means			
M	2	M	4.234	4.027	4.404	4.222	4.194	3.927	4.233	4.118
M	2	F	4.110	3.871	4.169	4.050	4.124	3.844	4.353	4.107
F	2	M	3.753	3.806	3.938	3.832	3.370	3.500	3.813	3.561
F	2	F	3.829	4.006	4.057	3.964	3.915	3.963	3.937	3.938
M	1	M	4.005	4.269	3.978	4.084	4.139	4.248	4.169	4.186
M	1	F	3.959	4.489	3.970	4.140	4.023	4.650	4.002	4.225
F	1	M	4.176	3.784	4.553	4.171	3.938	3.709	4.361	4.003
F	1	F	3.683	4.043	4.000	3.909	3.825	4.194	3.986	4.002
M	2	M or F	4.172	3.949	4.287	4.136	4.159	3.885	4.293	4.112
F	2	M or F	3.791	3.906	3.997	3.898	3.643	3.732	3.875	3.750
M	1	M or F	3.782	4.379	3.974	4.112	4.081	4.449	4.086	4.205
F	1	M or F	3.929	3.914	4.277	4.040	3.882	3.952	4.173	4.002
M or F	2	M	3.993	3.917	4.171	4.027	3.782	3.713	4.023	3.840
M or F	2	F	3.970	3.939	4.113	4.007	4.020	3.903	4.145	4.023
M or F	1	M	4.091	4.027	4.265	4.128	4.039	3.979	4.265	4.094
M or F	1	F	3.821	4.266	3.985	4.024	3.924	4.422	3.994	4.113
M	1 or 2	M	4.119	4.148	4.191	4.153	4.167	4.088	4.201	4.152
M	1 or 2	F	4.035	4.180	4.069	4.095	4.073	4.247	4.177	4.166
F	1 or 2	M	3.965	3.795	4.245	4.002	3.654	3.605	4.087	3.782
F	1 or 2	F	3.756	4.025	4.028	3.936	3.870	4.078	3.961	3.970
M	1 or 2	M or F	4.077	4.164	4.130	4.124	4.120	4.167	4.189	4.159
F	1 or 2	M or F	3.860	3.910	4.137	3.969	3.762	3.842	4.024	3.876
M or F	1 or 2	M	4.042	3.972	4.218	4.077	3.911	3.846	4.144	3.967
M or F	1 or 2	F	3.895	4.102	4.049	4.016	3.972	4.163	4.069	4.068
M or F	2	M or F	3.982	3.928	4.142	4.017	3.901	3.808	4.084	3.931
M or F	1	M or F	3.956	4.147	4.125	4.076	3.981	4.200	4.130	4.104
M or F	1 or 2	M or F	3.969	4.037	4.134	4.046	3.941	4.004	4.107	4.017

dence than parallel subjects whose sib was of their own sex (Tables 5, 10).

15. Girls with a brother differing two to four years in age exceeded girls with a sister in the rating they received on finality of decision, while sib's-sex-group differences among boys were insignificant. Children whose sib was of like sex and older by four to six years were rated less vacillating than comparable children whose sib was of unlike sex (Tables 5, 10).

16. The judgments for cheerfulness were more favorable for children whose sibling was dissimilar in sex than for those whose sib was of similar sex, when the sib disparity in age ranged from two to four years (Tables 4, 10).

TABLE 7

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: INDIRECTNESS OF RESPONSE TO FEAR AND FRUSTRATION AND SENSITIVENESS*

Subgroup			Indirectness of Response to Fear and Frustration				Sensitiveness*			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	4.276	4.124	4.398	4.266	4.190	3.959	3.991	4.047
M	2	F	3.885	4.069	4.328	4.094	3.854	3.665	3.699	3.740
F	2	M	3.386	3.586	3.931	3.634	4.035	4.216	4.093	4.114
F	2	F	3.815	4.333	3.761	3.970	4.323	3.699	4.288	4.104
M	1	M	3.992	4.566	4.114	4.224	4.015	4.009	4.401	4.142
M	1	F	4.061	4.469	4.283	4.271	3.700	3.723	4.233	3.885
F	1	M	3.699	3.746	4.118	3.854	4.151	4.192	4.323	4.222
F	1	F	4.161	3.604	3.931	3.898	3.998	4.120	3.791	3.970
M	2	M or F	4.080	4.097	4.363	4.180	4.022	3.812	3.845	3.893
F	2	M or F	3.600	3.959	3.846	3.802	4.179	3.958	4.190	4.109
M	1	M or F	4.026	4.518	4.198	4.247	3.858	3.866	4.317	4.013
F	1	M or F	3.930	3.675	4.024	3.876	4.074	4.156	4.057	4.096
M or F	2	M	3.831	3.855	4.164	3.950	4.113	4.087	4.042	4.080
M or F	2	F	3.850	4.201	4.044	4.032	4.089	3.682	3.994	3.922
M or F	1	M	3.845	4.156	4.116	4.039	4.083	4.100	4.362	4.182
M or F	1	F	4.111	4.037	4.107	4.085	3.849	3.992	4.012	3.927
M	1 or 2	M	4.134	4.345	4.256	4.245	4.103	3.984	4.196	4.094
M	1 or 2	F	3.973	4.269	4.305	4.182	3.777	3.694	3.966	3.812
F	1 or 2	M	3.542	3.666	4.024	3.744	4.093	4.204	4.208	4.168
F	1 or 2	F	3.988	3.968	3.846	3.934	4.160	3.910	4.040	4.037
M	1 or 2	M or F	4.053	4.307	4.281	4.214	3.940	3.839	4.081	3.953
F	1 or 2	M or F	3.765	3.837	3.935	3.846	4.127	4.057	4.124	4.102
M or F	1 or 2	M	3.838	4.005	4.140	3.994	4.098	4.094	4.202	4.131
M or F	1 or 2	F	3.980	4.119	4.076	4.058	3.969	3.802	4.003	3.924
M or F	2	M or F	3.840	4.028	4.104	3.991	4.101	3.885	4.018	4.001
M or F	1	M or F	3.978	4.096	4.111	4.062	3.966	4.011	4.187	4.055
M or F	1 or 2	M or F	3.909	4.062	4.108	4.026	4.033	3.948	4.102	4.028

Sib-Age-Difference-Group Disparities

17. Among first-borns there tended to be an increase with spacing in speed of emotional recovery, whereas in the case of second-borns there tended to be a decrease (Tables 2, 10).

18. Activeness in boys was, in the main, positively correlated with the degree of sib age disparity. No similar trend was observed for girls (Tables 8, 10).

TABLE 8

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: NERVOUS HABITS AND AMOUNT OF GROSS ACTIVITY*

Subgroup			Nervous Habits				Amount of Gross Activity*			
Subject		Sibling	Age Difference between Siblings in Months							
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total	7-24	25-48	49-72	Total
			Score Means				Score Means			
M	2	M	4.247	4.035	3.889	4.057	3.563	4.110	4.637	4.103
M	2	F	4.426	4.042	4.215	4.228	3.936	4.380	4.451	4.256
F	2	M	3.948	3.724	4.211	3.961	3.803	3.856	3.706	3.788
F	2	F	3.854	3.734	3.633	3.740	3.796	3.764	3.847	3.802
M	1	M	4.120	3.635	3.583	3.779	3.915	4.468	4.545	4.309
M	1	F	3.908	4.334	4.001	4.081	3.599	4.892	4.042	4.178
F	1	M	4.170	4.106	4.154	4.143	4.084	4.166	3.973	4.074
F	1	F	3.832	4.292	4.071	4.065	3.509	3.804	3.844	3.719
M	2	M or F	4.337	4.038	4.052	4.142	3.749	4.245	4.544	4.180
F	2	M or F	3.901	3.729	3.922	3.850	3.799	3.810	3.777	3.795
M	1	M or F	4.014	3.985	3.792	3.930	3.757	4.680	4.293	4.243
F	1	M or F	4.001	4.199	4.112	4.104	3.797	3.985	3.908	3.897
M or F	2	M	4.098	3.879	4.050	4.009	3.683	3.983	4.172	3.946
M or F	2	F	4.140	3.888	3.924	3.984	3.866	4.072	4.149	4.029
M or F	1	M	4.145	3.871	3.868	3.961	3.999	4.317	4.259	4.192
M or F	1	F	3.870	4.313	4.036	4.073	3.554	4.348	3.943	3.949
M	1 or 2	M	4.183	3.835	3.736	3.918	3.739	4.289	4.591	4.206
M	1 or 2	F	4.167	4.188	4.108	4.154	3.768	4.636	4.247	4.217
F	1 or 2	M	4.059	3.915	4.182	4.052	3.943	4.011	3.839	3.931
F	1 or 2	F	3.843	4.013	3.852	3.902	3.653	3.784	3.846	3.761
M	1 or 2	M or F	4.175	4.012	3.922	4.036	3.753	4.462	4.419	4.212
F	1 or 2	M or F	3.951	3.964	4.017	3.977	3.798	3.897	3.843	3.846
M or F	1 or 2	M	4.121	3.875	3.959	3.985	3.841	4.150	4.215	4.069
M or F	1 or 2	F	4.005	4.101	3.980	4.028	3.710	4.210	4.046	3.989
M or F	2	M or F	4.119	3.884	3.987	3.996	3.774	4.027	4.160	3.987
M or F	1	M or F	4.008	4.092	3.952	4.017	3.777	4.332	4.101	4.070
M or F	1 or 2	M or F	4.063	3.988	3.970	4.007	3.776	4.180	4.131	4.029

19. When the middle spacing is compared with the close, first-born boys at the former spacing scored the higher in readiness to anger, while second-born boys scored the lower. No significant spacing-group differences are to be noted in the case of the girls (Tables 4, 10).

20. As sib age disparity widened, there was an increase in self-confidence among boys with a brother, while among boys with a sister a peak occurred at the middle spacing and among F2F's a trough was exhibited (Tables 5, 10).

21. F1M's and F2M's manifested a maximum in finality of decision at the middle spacing, while F2F's at the same age-difference level evi-

TABLE 9

MEANS OF THE NORMALIZED RATINGS FOR THE VARIOUS SUBGROUPS: HEALTH

Subgroup			Health			
Subject		Sibling	Age Difference between Siblings in Months			
Sex	Ordinal Position	Sex	7-24	25-48	49-72	Total
			Score Means			
M	2	M	3.931	4.060	4.364	4.119
M	2	F	3.436	4.005	3.938	3.793
F	2	M	3.805	4.439	4.053	4.099
F	2	F	4.338	4.076	4.146	4.187
M	1	M	3.739	4.004	4.053	3.932
M	1	F	4.221	4.606	4.185	4.337
F	1	M	4.101	4.288	4.118	4.169
F	1	F	3.698	3.704	4.186	3.863
M	2	M or F	3.683	4.033	4.151	3.956
F	2	M or F	4.072	4.258	4.099	4.143
M	1	M or F	3.980	4.305	4.119	4.135
F	1	M or F	3.899	3.996	4.152	4.016
M or F	2	M	3.868	4.250	4.208	4.109
M or F	2	F	3.887	4.041	4.042	3.990
M or F	1	M	3.920	4.146	4.085	4.050
M or F	1	F	3.959	4.155	4.185	4.100
M	1 or 2	M	3.835	4.032	4.208	4.025
M	1 or 2	F	3.828	4.305	4.062	4.065
F	1 or 2	M	3.953	4.363	4.085	4.134
F	1 or 2	F	4.018	3.890	4.166	4.025
M	1 or 2	M or F	3.832	4.169	4.135	4.045
F	1 or 2	M or F	3.986	4.127	4.126	4.079
M or F	1 or 2	M	3.894	4.198	4.147	4.080
M or F	1 or 2	F	3.923	4.098	4.114	4.045
M or F	2	M or F	3.878	4.145	4.125	4.049
M or F	1	M or F	3.940	4.150	4.135	4.075
M or F	1 or 2	M or F	3.909	4.148	4.130	4.062

denced a low. Among M2M's there was a decrease in vacillation with spacing (Tables 5, 10).

22. Only at the two-to-four-year sib-age-disparity level did the older-sib group receive a higher mean rating on the trait, tendency to alibi, than did the younger-sib group (Tables 6, 10).

23. There tended to be an increase from the close to the middle spacing in the case of first-born boys in indirectness of response, while among F1F's the reverse trend was exhibited. Girls with a brother tended to manifest more indirection the more the sib differed in age (Tables 7, 10).

TABLE 10

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	Trait					
		Speed of Recovery from Emotional Disturbance			Social Apprehensiveness		
		Sum of Squares	Estimate of Variance	F	Sum of Squares	Estimate of Variance	F
Total	383	348.45			321.07		
Between groups	23	26.41	1.15		20.06	.87	
Within groups	360	322.04	.89		301.01	.84	
Between replications	15	12.24	.82		10.93	.73	
Residual	345	309.80	.90		290.08	.84	
Between variables							
Between sex groups	1	.18	.18		1.67	1.67	
Between sibling's-sex groups	1	.01	.01		1.58	1.58	
Between ordinal-position groups	1	1.98	1.98		1.30	1.30	
Between spacing groups	2	.50	.25		3.31	1.66	
First-order interactions							
Subject's sex--sibling's sex	1	.14	.14		.73	.73	
Subject's sex--ordinal position	1	.05	.05		.14	.14	
Subject's sex--spacing	2	.08	.04		2.90	1.45	
Sibling's sex--ordinal position	1	.36	.36		.48	.48	
Sibling's sex--spacing	2	.39	.19		1.51	.76	
Ordinal position--spacing	2	5.51	2.76	3.07	.04	.02	
Second-order interactions							
Subject's sex--sibling's sex--ordinal position	1	.29	.29		.54	.54	
Subject's sex--sibling's sex--spacing	2	5.49	2.74	3.06	3.36	1.68	
Subject's sex--ordinal position--spacing	2	3.52	1.76		.78	.39	
Sibling's sex--ordinal position--spacing	2	.01	.01		1.19	.60	
Third-order interaction							
Subject's sex--sibling's sex--ordinal position--spacing	2	7.91	3.96	4.41	.52	.26	

Sex-Group Differences

24. Boys with an older brother differing in age by less than two years were exceeded by girls with an older brother in speed of recovery of emotional poise, whereas the trend was reversed for first-borns. Girls with a brother two to four years younger were rated higher than boys with a brother similarly spaced, while among second-borns a parallel sex difference did not obtain. M2M(49-72)'s received a higher rating than F2M(49-72)'s (Tables 2, 10).

25. Estimates of the severity and number of nervous habits were higher for second-born boys than girls; whereas among first-borns the sex-difference trend was, if anything, antipodal (Tables 8, 10).

TABLE 10 (*continued*)

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	Cheerfulness			Tendency to Anger		
		Sum of Squares	Estimate of Variance	F	Sum of Squares	Estimate of Variance	F
Total	383	278.67			342.91		
Between groups	23	18.99	.83		27.67	1.20	
Within groups	360	259.68	.72		315.24	.88	
Between replications	15	9.62	.64		19.05	1.27	
Residual	345	250.06	.72		296.19	.86	
Between variables							
Between sex groups	1	.04	.04		6.19	6.19	7.21
Between sibling's-sex groups	1	.04	.04		.11	.11	
Between ordinal-position groups	1	2.25	2.25		.83	.83	
Between spacing groups	2	1.38	.69		.39	.20	
First-order interactions							
Subject's sex--sibling's sex	1	3.32	3.32	4.59	.26	.26	
Subject's sex--ordinal position	1	.29	.29		2.53	2.53	
Subject's sex--spacing	2	2.60	1.30		.90	.45	
Sibling's sex--ordinal position	1	.15	.15		.12	.12	
Sibling's sex--spacing	2	.15	.07		2.57	1.28	
Ordinal position--spacing	2	.64	.32		5.72	2.86	3.33
Second-order interactions							
Subject's sex--sibling's sex--ordinal position	1	.14	.14		1.14	1.14	
Subject's sex--sibling's sex--spacing	2	3.20	1.60		.22	.11	
Subject's sex--ordinal position--spacing	2	2.75	1.38		6.44	3.22	3.75
Sibling's sex--ordinal position--spacing	2	.63	.31		.02	.01	
Third-order interaction							
Subject's sex--sibling's sex--ordinal position--spacing	2	1.40	.70		.24	.12	

26. Boys were judged more active physically than girls only when the sib-age difference was more than two years (Tables 8, 10).

27. Boys, when the sib was female and younger by under four years, received a better rating in health than did girls with a younger sister, but when the younger sib was male, girls received a higher health assessment than boys, though not a significantly higher one. Among second-borns, F2F(7-24)'s were rated higher than M2F(7-24)'s (Tables 9, 10).

28. Second-born boys averaged higher in rating on readiness to anger than did girls; and the same trend, though not marked, appeared among first borns as well, when the sib was female. However, at the close spacing

TABLE 10 (*continued*)

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	Trait					
		Self-Confidence			Finality of Decision		
		Sum of Squares	Estimate of Variance	F	Sum of Squares	Estimate of Variance	F
Total	383	320.32			328.26		
Between groups	23	23.30	1.01		29.34	1.28	
Within groups	360	297.02	.83		298.92	.83	
Between replications	15	13.60	.91		21.22	1.41	
Residual	345	282.43	.82		277.70	.80	
Between variables							
Between sex groups	1	1.50	1.50		.00	.00	
Between sibling's-sex groups	1	.99	.99		2.33	2.33	
Between ordinal-position groups	1	3.81	3.81	4.64	2.04	2.04	
Between spacing groups	2	.72	.36		.22	.11	
First-order interactions							
Subject's sex--sibling's sex	1	.78	.78		1.53	1.53	
Subject's sex--ordinal position	1	.28	.28		1.94	1.94	
Subject's sex--spacing	2	3.91	1.95		2.39	1.19	
Sibling's sex--ordinal position	1	.93	.93		.01	.01	
Sibling's sex--spacing	2	.41	.21		5.38	2.69	3.34
Ordinal position--spacing	2	.05	.03		.74	.37	
Second-order interactions							
Subject's sex--sibling's sex--ordinal position	1	2.02	2.02		.70	.70	
Subject's sex--sibling's sex--spacing	2	5.57	2.79	3.39	7.40	3.70	4.59
Subject's sex--ordinal position--spacing	2	.23	.12		.38	.19	
Sibling's sex--ordinal position--spacing	2	1.59	.80		.25	.12	
Third-order interactions							
Subject's sex--sibling's sex--ordinal position--spacing	2	.50	.25		4.06	2.03	

the girl with a brother was judged higher than the boy with a brother (Tables 4, 10).

29. The rating given boys in social apprehensiveness was higher than that given girls, when the sib-age difference was under two years (Tables 3, 10).

30. When child and sib were near in age, girls were judged more self-confident than boys; but at the middle spacing the sex difference was reversed in the case of the children with a sister (Tables 5, 10).

31. Girls with a male sib were assessed less vacillating than boys when child and male sib differed by two to four years, but M2F's exceeded F2F's (Tables 5, 10).

TABLE 10 (*continued*)

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	Trait					
		Indirectness of Response to Fear and Frustration			Sensitiveness		
		Sum of Squares	Estimate of Variance	F	Sum of Squares	Estimate of Variance	F
Total	383	344.93			360.24		
Between groups	23	32.07	1.39		19.32	.84	
Within groups	360	312.86	.87		340.92	.95	
Between replications	15	10.30	.69		14.22	.95	
Residual	345	302.56	.88		326.70	.95	
Between variables							
Between sex groups	1	12.71	12.71	14.50	2.14	2.14	
Between sibling's-sex groups	1	.36	.36		4.45	4.45	4.70
Between ordinal-position groups	1	.45	.45		.19	.19	
Between spacing groups	2	2.73	1.37		1.81	.91	
First-order interactions							
Subject's sex--sibling's sex	1	1.80	1.80		.54	.54	
Subject's sex--ordinal position	1	.02	.02		.43	.43	
Subject's sex--spacing	2	.48	.24		.56	.28	
Sibling's sex--ordinal position	1	.04	.04		.30	.30	
Sibling's sex--spacing	2	.78	.39		.58	.29	
Ordinal position--spacing	2	.28	.14		1.63	.81	
Second-order interactions							
Subject's sex--sibling's sex--ordinal position	1	1.31	1.31		.51	.51	
Subject's sex--sibling's sex--spacing	2	3.12	1.56		.72	.36	
Subject's sex--ordinal position--spacing	2	5.38	2.69	3.07	2.69	1.35	
Sibling's sex--ordinal position--spacing	2	2.40	1.20		1.02	.51	
Third-order interaction							
Subject's sex--sibling's sex--ordinal position--spacing	2	.22	.11		1.74	.87	

32. M2M's at the over-four-year spacing were judged more final in decision than F2M's; but, when the sib was female, the trend was reversed (Tables 5, 10). Among first-borns at the same spacing level the estimate for M1F's exceeded that for F1F's.

33. Second-born girls at the close spacing were thought more cheerful than boys, whereas, at the two-to-four-year sib-age-difference level, those from opposite-sex pairs were rated more cheerful than those from same-sex pairs (Tables 4, 10).

34. Boys were rather consistently given higher ratings than girls in the traits, tendency to alibi and to respond indirectly to fear and frustration (Tables 6, 7, 10).

TABLE 10 (*continued*)

ANALYSIS OF THE VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	Trait					
		Tendency to Alibi			Health		
		Sum of Squares	Estimate of Variance	F	Sum of Squares	Estimate of Variance	F
Total	383	328.41			302.63		
Between groups	23	25.88	1.13		24.06	1.05	
Within groups	360	302.53	.84		278.56	.77	
Between replications	15	14.43	.96		9.67	.64	
Residual	345	288.10	.84		268.89	.78	
Between variables							
Between sex groups	1	7.33	7.33	8.78	.16	.16	
Between sibling's-sex groups	1	.74	.74		.17	.17	
Between ordinal-position groups	1	2.43	2.43		.03	.03	
Between spacing groups	2	1.89	.94		4.33	2.17	
First-order interactions							
Subject's sex--sibling's sex	1	.84	.84		.44	.44	
Subject's sex--ordinal position	1	.72	.72		2.05	2.05	
Subject's sex--spacing	2	.64	.32		.61	.31	
Sibling's sex--ordinal position	1	.88	.88		.58	.58	
Sibling's sex--spacing	2	2.01	1.00		.36	.18	
Ordinal position--spacing	2	1.79	.90		.10	.05	
Second-order interactions							
Subject's sex--sibling's sex--ordinal position	1	.98	.98		7.23	7.23	9.28
Subject's sex--sibling's sex--spacing	2	1.01	.50		3.99	2.00	
Subject's sex--ordinal position--spacing	2	2.88	1.44		1.71	.86	
Sibling's sex--ordinal position--spacing	2	1.59	.80		.25	.12	
Third-order interaction							
Subject's sex--sibling's sex--ordinal position--spacing	2	.16	.08		2.05	1.02	

Discussion of Group Results

M2M.—The boy with an older brother, when compared with the boy with a younger one, if the sib differed in age by less than two years, was judged as more ready to anger and less apprehensive in physical activities (not quite significant). At the close spacing the former doubtless competes vigorously, has to keep on his toes to defend himself, and is encouraged to fight back. He also, probably, since he is a second-born, has been warned less of the dangers of the physical world and has had his older sib's example to encourage him in the effort to master body control and the physical environment. When the age gap between the members of MM pairs was two to four years and M2M's, hence, were much outclassed

TABLE 10 (*continued*)

ANALYSIS OF VARIANCE FOR VARIOUS TRAITS

Source of Variance	df	Nervous Habits: Sum of Squares	Nervous Habits: Estimate of Variance	Nervous Habits: F	Amount of Gross Activity: Sum of Squares	Amount of Gross Activity: Estimate of Variance	Amount of Gross Activity: F
Total	383	337.80			332.57		
Between groups	23	20.03	.87		49.04	2.13	
Within groups	360	317.76	.88		283.53	.79	
Between replications	15	13.20	.88		8.62	.57	
Residual	345	304.57	.88		274.90	.80	
Between variables							
Between sex groups	1	.17	.17		13.26	13.26	16.64
Between sibling's-sex groups	1	.14	.14		.71	.71	
Between ordinal-position groups	1	.02	.02		.56	.56	
Between spacing groups	2	.68	.34		12.11	6.06	7.60
First-order interactions							
Subject's sex--sibling's sex	1	2.97	2.97		.89	.89	
Subject's sex--ordinal position	1	6.00	6.00	6.80	.02	.02	
Subject's sex--spacing	2	1.73	.87		8.32	4.16	5.22
Sibling's sex--ordinal position	1	.38	.38		2.75	2.75	
Sibling's sex--spacing	2	1.71	.85		.81	.41	
Ordinal position--spacing	2	1.59	.80		2.18	1.09	
Second-order interactions							
Subject's sex--sibling's sex--ordinal position	1	.04	.04		.07	.07	
Subject's sex--sibling's sex--spacing	2	1.48	.74		3.88	1.94	
Subject's sex--ordinal position--spacing	2	.20	.10		1.86	.93	
Sibling's sex--ordinal position--spacing	2	2.41	1.20		1.16	.58	
Third-order interaction							
Subject's sex--sibling's sex--ordinal position--spacing	2	.53	.26		.46	.23	

in ability by their sibs, the former could not have had much expectation of success for themselves and must have functioned chiefly as hangers-on in the brothers' play groups. In spite of this, M2M's at this spacing were judged, in comparison with M1M's, to recover more readily from emotional upset, to be less intense, more cheerful, less likely to anger, and less inclined to be indirect in response to fear and frustration; i.e., M2M's, relative to M1M's at the middle spacing, seem more poised. It is relevant to recall that M2M's not only did not experience displacement as did M1M's, but this displacement for the latter came at a very critical time in their dependency development and probably resulted in their being rather anxious, hostile, and competitive (see References 12, 18, 19). M2M(25-48)'s, on the other hand, doubtless received much protection from their mothers and possibly a more consistent level of appreciation than M1M's (see Reference 12). Congruent with the hypothesis that M2M's were babied and indulged by the parents though chastened by the sib is the indication in Reference (9) that M2M(25-48)'s made many bids for help and attention from the teacher, who is a sort of mother figure.

Moreover, the brother's type (M1M) showed, contrary to M2M's relative poise, marked signs of stress and jealousy (10) at this two-to-four-year spacing—another fact which lends support to our hypothesis that M2M's probably were having it well. In fact, they seemed to exhibit better adjustment as the sib age gap widened; i.e., the ratings show that M2M's at the widest spacing, in contrast to the close, received a significantly lower rating in vacillation, in readiness to anger, in social apprehensiveness and a higher one in cheerfulness, self-confidence, and activeness. The ratings given the M2M groups also reveal, as the sib age gap broadened, a marked, but not quite significant, change in the positive direction in the case of the traits of health and speed of recovery from emotional upset, as well as a drop in nervous habits. The view that the greater the sib age difference the better adjusted the M2M's were is further bolstered by data we have (10) which indicate that at the widest spacing the boys with an older brother were more sociable and expansive than the boys whose brother was less than two years older. We think the greater sociableness and better adjustment of the M2M's at the wide spacing occurs because, when the sib is very much older, the younger child tends to have more companions of his own choosing, as well as more protection and relative appreciation from his mother—possibly even from his sib—while, at the same time, good vigorous reinforcement for his male sex identification from both his brother and father. Our findings (10) suggest, too, better adjustment of the brothers of the M2M's when the sib spacing was wide than when close—a condition that should make the adjustment of the younger child less stressful. The fact of the better health of M2M's the wider the sib spacing, it is our belief, stems from lesser exposure from the sib. The latter by the time our subjects entered school had probably already had many of the children's diseases and by virtue of his age had also become relatively resistant to colds. Hence he would not have been much of a source of contagion to his five- or six-year-old sib, our subject.

Compared with the boy with an older sister, M2M's seem, at the wide spacing at least, more expansive and poised. M2M's received at the over-four-year sib-age-difference level a lower mean rating in social apprehensiveness and nervous habits and a higher rating in finality of decision, self-confidence, and recovery from upsets. We shall comment on this difference when we discuss the case of M2F in the next section.

M2F.—The boy with an older sister, as stated above, presents a somewhat different picture from the boy with an older brother, at least when the sib-age-difference is over four years. The apparent less expansive adjustment of M2F's at the wider spacings, compared with M2M's, may have been due to: (1) sex role conflict and greater exposure to females on the part of the former (M2F's models were chiefly females and the father, if Freudian theory is correct, may have favored the sib); (2) the pressure from the former's rivalrous, dominant, and jealous older sister (see References 9 and 10 and data on F1M for evidence relative to the sister's type); (3) the babying and indulgence the former doubtless received from his mother because he was the younger of her children by much, a male (see Sears *et al.,* 17), and different in sex from the sib. Other findings we have (9, 10) show our M2F's to have been at the widest spacing both less friendly to the teacher and peers and more hostile than our M2M's. The former, according to our interview data, were also particularly given to quarreling with their sibs. M2F's, the wider the spacing, showed less good emotional balance, more gross physical activity, better health, but also more tendency to alibi and to respond indirectly to fear and frustration. The constellation of traits the M2F's exhibited, especially as the age gap between them and their sibs expanded, was that characteristic of somewhat dependent, passive, spoiled children. It is even possible that M2F's health, which was reported less good than the M2M's, was actually good but the former were kept out of school on less provocation by their indulgent mothers. However, M2M's may have been exposed to more diseases by their sibs in their preschool years than were M2F's and hence may have entered school with more immunities. It is worthy of note that M1M's health in the early school years was relatively poor, while F1M's was relatively good. These two groups are the sib types for the M2M's and M2F's.

When M2F's are compared with M1F's, their counterparts among first-borns, we find at the close spacing the two groups differed little, the former having been judged to recover more speedily from disturbance, to have more nervous habits and possibly poorer health. At the widest spacing, where, as was suggested earlier, M2F's seemed somewhat spoiled and ambivalent, they were judged to be less sensitive and more vacillating. M2F(25-48)'s, relative to M1F(25-48)'s, were rated less confident, active, healthy, less given to anger, to alibiing, and to projecting blame, and less intense, excitable, and moody. The former seem to show more poise but also a less dominant attitude than the much threatened but expansive M1F's. We suspect the differences described derive from M2F's being more indulged and having experienced less stress than the M1F's, who were displaced at a critical period in their dependency development [see similar comment above regarding M1M(25-48)'s]. M2F(25-48)'s sib type, like M2M(25-48)'s, seemed goaded by jealousy and was dominant and expansive. Our data give the impression, then, that having an older, in contrast to a younger, sister was probably slightly more stressful for the boy when the sib spacing was very close but clearly less stressful when the spacing was two to four years. The finding may obtain only at the age

Bureau Edn. Res.
Dated
No.

level with which we are dealing, for never again will the sib be so attractive. Ordinal position was correlated with a few of the variables in this study, however, when the sib age gap was four to six years; but other data we have (10) indicate M2F(49-72)'s were more withdrawn and less competitive than M1F (49-72)'s.

F2M.—The girls with an older brother, like the boys with an older sister, showed signs of greater passivity and of increased stress when the sib deviated from them most in age—i.e., the ratings indicate that there occurred, as the sib age difference broadened, a decrease in cheerfulness, self-confidence, finality of decision, and speed of recovery from upset, as well as an increase in nervous habits, indirectness of response to frustration, tendency to alibi, resistance, jealousy, criticalness, tendency to bid for adult attention, and to be upset readily by defeat (for a discussion of the latter traits see References 9, 10, 11). Also the wider the sib age difference the less friendly F2M's were to adults, although they increased their bids for attention as the disparity in years between them and their brothers expanded; i.e., as the brothers' dominance increased, his chastening could be less well dealt with, and the advantages he had became ever more apparent. F2M's mothers, moreover, if the theory of cross-sex alliances in the family is correct, possibly may have been more devoted to the sib, who was their first and long their only love. We note in this connection that, when we interviewed the children, they reported with *decreasing* frequency as the sib age difference expanded that they thought the mother sided with them when they quarreled with the sib. F2M's at the wide spacing were, according to another division of our investigation (11), rated low in responsibleness and tenacity—an observation that is congruent with the hunch that these children were frustrated and felt they were competing at a disadvantage. F2M's at the spacings under four years doubtless received a good workout from their brothers but were able to defend themselves to some extent. However, at the widest spacing, when F2M's were clearly no competitors for their sibs in strength or ability and were probably shunned by their brothers and the latters' friends, they might well have felt inadequate and poorly compensated for their lowly status by the protection and indulgence that they, as the youngest in the family, received.

When the girls with an older brother are set in contrast to the girls with a brother correspondingly younger, the picture for the former at the under-two-year sib-age-difference level is perhaps one of better adjustment, as was the case also for M2F's in relation to M1F's. F2M(7-24)'s, in contrast to F1M(7-24)'s, for instance, received ratings indicating a tendency to more rapid recovery from emotional upset, greater cheerfulness, less self-confidence, but also less alibiing and projection of blame. The former did not experience displacement, profited by the parents' experience with the first child, and had to learn to compete with the sib from the first. These are all influences likely to favor good adjustment. F2M's at the wide spacing, on the other hand, seem relatively more constricted than F1M's, a relation similar to that which held between M2F's and M1F's. The ratings indicate significantly less rapid recovery of emotional balance, less self-confidence, more social apprehensiveness, less tendency to anger, to alibi, to project blame, as well as less sensitiveness. F2M(49-72)'s were also judged less affectionate, less demanding of attention, and less responsive to adults (9). F1M(49-72)'s were long only children and as such were probably given more stimulation by adults, more appreciation and responsibility than were F2M(49-72)'s. The formers' defenses were those more effective with adults, e.g., alibiing; whereas the defenses of the latter were more those of the less able, e.g., silence and withdrawal. It is interesting, apropos of our hypothesis that F2M(49-72)'s were both babied and chastened more than F1M(49-72)'s, that the former were rated high in jealousy, low in responsibleness, and high in dawdling (see References 10, 11). These are lineaments of the dependent child.

Comparing F2M's and F2F's to get leads as to the effect of the sex of the sib, when age difference is constant, we find, in contrast to the relation between M2M's and M2F's, few significant differences. If we pay attention to the middle spacing, where disparities are most numerous, we note F2M's were rated significantly less apprehensive socially, less indirect in response, more self-confident, but also more sensitive than F2F's. Other findings we have indicate the former were less vacillating, less critical, more competitive, friendly, popular, and better leaders (10). At the middle spacing, F2M's, we think, are beginning to find themselves not too welcome in their brothers' play groups and are seeking associates outside the home, whereas F2F's are still functioning as satellites of their older sisters. At the four-to-six-year age-difference level, where, whatever the sex of the sibs, association with them is limited, the only significant disparity between F2M's and F2F's occurred in the case of nervous habits, F2M's being thought to have the more. When, on the other hand, the sibs were nearly the same age, F2M's were judged to alibi less, to be less indirect, and to have less good health. In the main, then, F2M's at the closer spacings seem somewhat more direct or less shy than F2F's. They were, according to other findings of ours (9, 10), less inclined to bid for adult attention, to insist on their rights, to tattle and were more popular. F2M's at the closer spacings had more friends of their own than did the F2F's. This may have been responsible for their greater social effectiveness, as well as account for their less good health. A similar type of relation in the case of health and the sib's-sex variable was described earlier when M2F's were contrasted with M2M's.

F2F.—The girls with a sister older by under two years, when compared with those with a sister younger by less than two years, like the other second-borns in relation to the first-borns at this spacing, seem to have some advantageous traits. They were judged less moody, less fearful of physical activities (a similar pattern was noted when M2M's were compared with M1M's), less vacillating, and possessed of better health. Probably F2F(7-24)'s had relatively more of the children's diseases before entering school because of exposure from their older sibs, with whom they played constantly, and thus had more immunity on entering school. F1F's, on the other hand, having associated with a lesser range of children in their preschool years than F2F's and, hence, having probably had fewer children's diseases, doubtless succumbed readily to the infections to which they were exposed at school.

F2F(7-24)'s may have been somewhat better adjusted than F1F(7-24)'s because they had no scars from a displacement experience, had probably been dealt with in a more relaxed fashion than the latter by virtue of being second-borns, had virtually no sex identification problems, and had had practice in associating with children from the beginning. In addition, they had a reasonable chance of holding their own in competition with their sib, since the latter was only slightly older. However, at the middle spacing, when F2F's were clearly outclassed by their sibs and were hangers-on in the sibs' play groups, they scored significantly lower in self-confidence and higher in indirection than did F1F's who, though challenged by the sister, nevertheless felt themselves to be in the driver's seat. At the widest spacing, the F2F's were rated higher in sensitiveness and finality and lower in nervous habits and anger than F1F's. Other evidence we have indicates F2F(49-72)'s were the more jealous and resistant (9, 10). The former, we believe, are still satellites of the sister, though probably not so much so as at the middle spacing, and adopt frequently a pattern of defense favorite with the less able, namely, resistance. This tendency to resist may lie back also of F2F's "decisiveness." If we look at the spacing-group differences among F2F's, we see evidence at the closer spacings of a strong drive to keep up with the sib but evidence of greater passivity, resignation, and dependency when the sib is over four years older. The most stress is suggested at the middle spacing. Comparing, for instance, the children at the under-two with those at the two-to-four-year sib-age-difference level, we note the direction of change was toward greater apprehensiveness, less self-confidence, and less finality. This middle spacing is the level at which, as mentioned above, we believe, F2F's are chiefly hangers-on in the sisters' play groups. F2M's, in contrast, at the middle spacing, because of sex distance (7), probably have already begun to have friends of their own sex, age, and choosing. This may account for the favorable change among F2M's from the close to the middle spacing, while among F2F's the shift seems in the direction of greater tension or dependence. The reversing of the direction of change in the F2F's, relative to the F2M's, it seems reasonable to infer, is due in part to the greater age difference necessary to give those children whose sib is older and similar in sex, as compared with those whose sib is older and opposite in sex, relative independence from the sib and its companions. Also at the wide spacing, when indulgence of the F2F's might be great because of their being so much the younger sib, we should expect less difference in the treatment meted out to the members of F2F and F2M pairs. When the age difference between the F2F's and their sisters is wide enough so that the former do have associates of their own, our data suggest their emotional-social adjustment takes an upturn. For example, F2F(49-72)'s, relative to F2F's at the two-to-four-year spacing, were judged more self-confident, less vacillating, more cheerful, less indirect, and less moody.

M1M.—M1M's present a rather interesting pattern which seems repeated in some ways in the other groups of first-borns. The ratings of M1M's at the close spacing suggest they were relatively passive and dependent; i.e., low in initiative and not very friendly to their classmates (11, 10) but friendly to adults (9) and indirect. When, however, the sibs of the M1M's were two to four years younger, the latter showed more dominance and were more actively defensive, less adult-oriented, and more resistant. The change, when the two spacing groups are compared, is in the direction of greater emotional intensity, moodiness, indirection, anger, activeness, fewer nervous habits, slower speed of recovery from emotional disturbance, less cheerfulness, and less apprehensiveness in physical activities. It is our hunch that M1M(25-48)'s showed the great drive and stress they did because the sib is at that adorable one-to-four-year age when it receives a great deal of affectionate attention and appreciation from people generally. It gets around on its own, too, and seriously interferes with the older child's activities and possessions. M1M(25-48)'s, it is true, are abler than the sib and can usually defend themselves physically but may experience much frustration partly because they are discouraged by their parents from attacking the sib directly. Some might stress the importance of the picture of the time of arrival of the sib (19) and of the jealousy experiences, taking these to occur in the case of M1M(25-48)'s at a particularly critical period in the latter's dependency learning and hence to cause much stress (see References 12, 18, 19). The attitude of status concern, dominance, alertness in the defense of self interests, built up in the home because of the sib situation, it would be reasonable to expect M1M's to carry over to their peers. Whatever the cause, M1M(25-48)'s dared to lash out. But M1M-(7-24)'s clung more to adults and possibly even to the sibs. Since they had a sib companion near their own age almost from the first and probably were the recipients of treatment from the parents rather similar to that given the brother, this might have caused them to find the latter's companionship sufficiently satisfying to curb somewhat the desire to seek companionship out of the home. M1M(7-24)'s were rated relatively low in friendliness to children, for instance, but in the interview rarely said they thought they could dispense with the sib. When M1M's sib, on the other hand, is virtually an infant (wide spacing) instead of more nearly his age, we should expect the latter to fire the former's drive little and to interfere inconsequentially with his activities. Consistent with this hunch, we observe that M1M's at the widest spacing, relative to M1M's at the close, were judged to have fewer nervous habits, to be the more cheerful, less moody, more active, and less intense. And, if M1M's whose sibs are younger by four or more years are compared with M1M's at the two-to-four-year spacing, it is to be noted that the former were rated as recovering emotional poise more readily, as more cheerful, less indirect, more apprehensive in physical matters, and less moody. M1M's, it is our belief, tend at all spacings to identify some with the brother and to have their sex identification reinforced by the latter.

In contrast to M1M's, M1F's do not have their sex identification reinforced by the sib and, in addition, may have been deterred more, because the sib was a girl, from attacking her directly, when self interests were threatened. However, in spite of the security-threatening considerations just mentioned, M1M(25-48)'s, we think, have more direct problems with the sib than M1F(25-48)'s do because of more frequent association with the sib. This may account for M1M(25-48)'s being judged the less confident, cheerful, and active. M1M's speed of recovery also was rated less at the middle spacing than was that of the dominant M1F's, and the former were thought to have fewer nervous habits and less good health, as well as to be less self-confident, less active, and less cheerful. At the close spacing, on the other hand, M1M's were rated higher in the speed with which they regained emotional balance after upset but at the same time were thought more apprehensive than M1F's. The pattern of less good health at the closer spacings in the M1M's than among the M1F's has some consistency and may stem from the former having had a less wide range of associates in their preschool years, less exposure to disease, and hence, less occasion to acquire immunities before entering school.

If we pay attention to ordinal-position differences among the MM's, we observe that, when the sibs differed little in age, M1M's were thought to show more constriction than the M2M's. For instance, the former were judged to anger less, to be less excitable, and to be more apprehensive. The former, we think, have been discouraged by the parents from overt expressions of anger and have been contaminated by the parents' anxiety and concern, while M2M's have probably been encouraged to defend themselves, to try to keep up with their older brothers, and to be less apprehensive about physical activities. At the middle spacing M1M's show clearly more tension than M2M's; i.e., the former were rated more intense, moody, less cheerful, more given to anger, slower to recover from emotional disturbances, and possessed of more nervous habits. At the wide spacing, when both groups of boys probably had a somewhat similar amount of parental attention, M1M's differed from M2M's only in sensitiveness and physical apprehensiveness.

M1F.—Boys with a younger sister show some of the same pattern details as far as spacing-group differences are concerned as do boys with a younger brother. Marked change, for instance, in the direction of greater dominance and active defensiveness occurs between boys with a sister close in age and those with the sister two to four years younger. M1F(25-48)'s, compared with M1F(7-24)'s, were judged to have more nervous habits, to be more active, sensitive, readily angered, self-confident, excitable, intense, to give more alibis, more frequently to project blame, and to be lower in social apprehensiveness. However, the boys with a sister four to six years their junior were thought less on the defensive and less dominant than those with a sister two to four years younger, i.e., as offering fewer alibis and projecting blame less, as less intense, excitable and active, and as more apprehensive in physical activities. The former do not have in the baby sister much of a direct competitor and the parents' behavior toward the baby is, we think, likely to be interpreted as less biased. Moreover, M1F(49-72)'s sibs arrived after they had developed relatively mature social interests—a fact that should make the reduction in parent attention less of a blow for them.

A look at the ordinal-position differences among the M1F's discloses that, at the close spacing, M1F's, compared with the M2F's, scored lower on nervous habits and in speed of recovery from emotional upset, a rather general characteristic of first-borns when the sib is near in age. Compared with the boys with a sister their senior by two to four years, M1F(25-48)'s seem more actively defensive. As stated in the previous section, the ratings given the former indicate that they were more active, possessed of better health, angered more readily, were more excitable, moody, intense, as well as inclined to project blame and alibi more. This tendency to blame-avoidance of first-borns, Sanford *et al.* (15) also noted. In the previous section we explained our opinion that this pattern results from the displacement experience at a critical time, M1F's sex identification conflict, and the great attractiveness of the sib at the ages one to four years. Although at the over-four-year spacing the ordinal-position differences are less marked, M1F's were judged more sensitive and less vacillating than M2F's. Other findings we have (10) show the former at the wide spacing to have been more gregarious, competitive, better leaders, and given less to teasing. On the whole, it looks as if they were somewhat better adjusted than their counterparts among second-borns.

To recapitulate briefly what was elaborated in the previous section, we note, if we contrast M1F's with M1M's, that when the sib was close in age the two groups differed little. M1F's were judged to recover poise less readily, but to be less apprehensive about physical activities and to have better health. They were also thought to be less friendly to their teachers and peers, more critical and insistent on rights (9, 10); i.e., they seemed under more social tension. At the middle spacing, while the tension signs are very apparent, M1F's appear to be more expansive. Relative to the boys with a younger *brother*, M1F's at the middle spacing were assessed as more healthy, self-confident, cheerful, given to alibiing, active, recovering emotional poise more speedily, and as having more nervous habits. They were also judged more jealous, exhibitionistic, competitive, popular, and better leaders, as well as more friendly to adults and children, inclined to bid for help and attention, and to tattle (9, 10). M1F's at the middle spacing probably have less of a *direct* competitor in the sib than do M1M's, but are much upset by the amount of adult attention the sister receives, this sib of a *different* sex! M1F's, it appears, are much activated by sex rivalry. The companionship status of the sibs of the M1F's, we think, changes most between the close and middle spacing (it is reduced), whereas for M1M's it should be noted, the change is greatest beween the middle and wide spacing. At the latter, where the sib and his sex are less in evidence, M1F's differed little from M1M's, being thought somewhat less active, more moody, and possessed of more nervous habits.

F1M.—The girls with a younger brother, like the M1M's, seem to have been spurred by the brother; but the sib's relative age appears to have made less difference than it did in the case of the latter group. Does the girl have a sufficiently low jealousy limen so that the *fact* of the sib rather than its *character* is the important variable? F1M's were rated very jealous, exhibitionistic, quarrelsome, competitive, insistent on their rights, as well as friendly to their peers and to adults (9, 10). At all spacings the F1M's were adult-centered but at the middle spacing they were somewhat less responsive to adults than at the close or wide, a fact that suggests the middle spacing for the girls, as for the boys previously discussed, was the most difficult one.

If we turn to spacing effects upon the variables on which this study is focused, we find that F1M(7-24)'s were judged more vacillating and less poised than F1M(25-48)'s. The former, relative to F1M(49-72)'s, who had the most association with adults, were thought to recover less quickly from upset, to alibi less, and to be less indirect in response to frustration. The assessments of the F1M(25-48)'s, in contrast with those for the F1M(49-72)'s, indicate the former were less moody, less vacillating, and less given to projecting blame, to indirection and to alibiing. Apparently F1M's, when the sib is an infant, are taxed little by him directly but suffer or grieve much because of the affectionate attentions the baby receives. Forty per cent of the F1M's at the wide spacing said they wished they could change places with the baby, whereas at the close spacing only seven per cent desired to make this exchange with the sib.

Compared with the girls with a younger sister, F1M's, when the sib was close in age, were judged to anger more readily, to be more active, more intense, to recover less quickly from disturbance, to be less given to indirect responses to fear and frustration, and more likely to project blame. In other words, having a male sib seems to make the girl more direct and vigorous in defense. In keeping with this picture, F1M's at the middle spacing were rated the less vacillating, more cheerful, and less inclined to alibi; and at the widest spacing, as less apprehensive socially, less intense, and more given to projecting blame. The F1M's, like the M1F's, who were also members of mixed-sex sib pairs, were much stimulated and their dominance, we think, was motivated in part by sex rivalry. But the F1M's may also have had less contact with their sibs than did the F1F's and have sought, to a greater extent, companionship with children out of the home. If this is the case, the former might be expected to be more dominant in their social relations. It is significant in relation to the question of the associates the child has out of the home that the F1M's expressed, with a relatively high frequency, a preference to play with some child other than the sib. It is also interesting that the sibs, if the M2F type is any indication of what the former were like, were probably rather docile, mother-centered, mother-indulged, and withdrawn, although quarreling much with their dominant older sisters (9, 10).

F1F.—The girls with a younger sister showed the expected signs of inflation in tension as the sib spacing increased from under two years to two to four years, but, of course, not so great a change occurred as did among the males. Girls with a sister two to four years their junior, as compared with those whose sib was less than two years younger, received a higher mean rating on nervous habits and tendency to project blame and a lower rating on cheerfulness and indirection. The reader will remember we have already suggested why there may be an increase in tension at the middle spacing and have offered some hypotheses as to why the girls show less increase than the boys. This discussion we shall not repeat here. Among F1F's, changing in the direction of increase over the whole range as the sib age difference widened were the traits, health, excitability, and intensity. Moodiness, on the other hand, decreased. The waxing in excitability, intensity, and anger as the spacing increased may be due in part to F1F's decreasing experience with children. F1F(49-72)'s may have been very excited when playing with peers at school because this was a relatively infrequent happening for them. F1F(49-72)'s were long only children. They probably were contaminated by the anxiety their parents had relative to them because they were the parents' first contact with children, and they were doubtless subjected to this contamination in a less dilute form than were those whose sibs arrived earlier in their lives.

For a discussion of what in F1F's behavior may be associated with the sib's-sex and ordinal-position variables, the reader is referred to the sections describing the F1M's and the F2F's.

Summary Comments

Although no detailed summary of the interrelations of all of the variables with which we were concerned would be feasible (for this, the reader is referred to pages 397 to 409), let us here comment briefly on a few of the major trends in our data. It is clear that birth-order differences which obtain at the under-two-year spacing may not obtain at the wider spacings. Ordinal-position-group differences at the close spacing, we think, reflect to a relatively greater extent direct sib interaction effects, whereas at the wider spacings the differences are increasingly expressions of child-parent relations, conditioned, to be sure, by the fact of the sib and his relation to the parents.

First-borns at the close spacing were judged to recover less readily from upsets as well as to anger less than second-borns. The former have more status to defend than the latter but are handicapped in a direct struggle with the sib by parental restraints. Hence the former are likely to be more concerned over their defeats. Also first-borns, we think, have stronger superegos. Their parents, unaccustomed to children when their first arrived, have expected more of them than is usual for later children. Second-borns, who are encouraged to defend themselves against the sib, are, on the other hand, less hesitant to express their anger and to attack directly than are first-borns.

To our surprise (15), first-borns were rated the more self-confident, a condition we now think may have stemmed from their being the abler members of the sib pairs. The self-confidence ratings, however, were based on school behavior. Since first-borns tend to be more adult-centered or oriented, the latter attitude may appear to the teacher as confidence. Indeed, in a school situation over which an adult presides, first-borns may be more self-confident than second-borns.

We had also expected first-borns generally to show more nervous habits than second-borns because a higher standard of impulse control is expected of them by adults (15). The "nervous habits," we thought, might reflect the drainage of tension. The fact that second-born males exhibited more nervous habits than did first-borns has made us reconsider. The second-borns *also* are curbed and frustrated but for a different reason, namely, their ineffectiveness in defending themselves against the older sib. This state of affairs may generate much tension. However, as explained earlier, we are not certain that the so-called nervous habits are reflections chiefly of emotional-social disturbance. We have the hunch that local irritations often are the instigators of what we name nervous habits; e.g., dry corneas, tight clothing, caked mucous in the nose, rough skin, hang-nails, skinned knuckles, or callouses. In the light of this surmise, we might hypothesize that second-born males acquire more local irritations in their efforts to keep up with the older sib and through these bruisings *et cetera,* more nervous habits. Since girls are less active and less given to anger, ordinal-position differences in nervous habits among them may reflect more the operation of their superegos than is the case for boys. If this is the situation, first-born girls should exhibit more nervous habits than second-born, which according to our findings, they do.

Another ordinal-position difference we find interesting is that among first-borns the children from opposite-sex sib pairs tended to have better health than those from same-sex pairs, while this pattern was reversed among second-borns. Contrary to the findings of Kingsley and Reynolds (6), we noted no simple ordinal-position differences in health. If our measure of health is a valid one, we think the above finding may mean the following: First-borns whose sibs are opposite in sex are more likely to seek associates of their own sex and age out of the home and thus are exposed to more contagion. They may, hence, acquire many immunities before entering school and thus succumb less frequently to disease in the early grades in school than do the previously less exposed children from same-sex pairs. The pattern of health among first-borns we would expect to produce the pattern of health we observe among second-borns. If the first-born in his early school years has a period of good health, then the second-born sib should be less exposed to disease and hence have less opportunity to acquire immunities, a condition that should result in less good health for him in the early grades in school. We think it possible, too, that second-borns from opposite-sex pairs, in contrast to those from same-sex pairs, are exposed more at school age to a wider range of children. Because of the phenomenon of sex distance (8), they may be cut off some from the friends of their school-age sib and this encourages them to seek associates of their own. This results in their having direct or indirect contact with more children and hence there is a greater likelihood of their exposure to contagion than is the case in the second-borns whose sib is of the same sex and whose playmates hence overlap much with those of the sib. When the sib is an infant or very much older than the child, then its sex we should expect to have little effect on the latter's associates and hence on their health. Our finding of a lack at the wide spacing of significant sib's-sex-group differences in health is congruent with the expectation stated.

If we turn to our data bearing on the effect of the sib's-sex variable, we note more signs of stimulation or stress among the members of pairs whose sib was different in sex than among members of sib pairs similar in sex, though the spacing and ordinal-position variables often interact significantly with the sib's-sex variable (see also References 9, 10, 11). The interaction of ordinal position, spacing, child's-sex, and sib's-sex variables in the case of health, we have already discussed. We observe also that the girl with the younger brother was noticeably on the *qui vive*. The most conspicuous sib's-sex group differences seemed to occur at the two-to-four-year sib-age-difference level. At this level the children whose sib was of a different sex from themselves seemed more self-confident, more cheerful, active, healthy, less vacillating, and more inclined to recover poise readily than the children whose sib was like them in sex. Apparently sex rivalry is stimulating. Or sib interests overlap somewhat less when the children differ in sex and direct sib conflict hence is less (7). In addition children from opposite-sex sib pairs may have a greater range of child companions and play less frequently, if second-borns, the role of satellites in their sibs' play groups.

It is also worthy of note among sib's-sex relations that children with a brother were judged less sensitive than those with a sister. Since the culture expects the male to take it on the chin, it may be this attitude is communicated directly or indirectly by the male to his sib. It is relevant in this connection that, in the main, our male subjects were rated less sensitive than the females, the p for the sex-group difference falling only slightly short of 5 per cent (Tables 7, 10).

Since the spacing variable was related in a very complex fashion to our other variables, we shall attempt no description here of the web of relations. Let us merely say that spacing differences seem greater in the case of the males than females. The two-to-four-year spacing, especially for the first-borns, appeared to be a stimulating and/or stressful one, though the meaning of the indications is not so clear as in other data we have reported (9, 10). Confidence, emotional intensity, excitability, moodiness, anger, decisiveness, alibiing, projecting of blame, and indirection, all tended to be high among first-born males at the middle spacing, as compared with those at the close spacing. The boys seemed stung to the quick by the attentions their one-to-four-year-old sibs received, particularly if the sibs were sisters. When the sibs are very near in age, they get more nearly similar treatment, a fact which no doubt accounts in part for the lesser stress at the close spacing.

In the main, gross activity and possibly health were positively correlated with spacing. We suspect the more the sibs differ in age the less they expose each other to diseases, or the narrower their range of child associates and the more time they spend with adults. These considerations may account for the positive relation between health and physical activity and spacing.

Simple sex differences occurred rarely. Boys generally were judged more active than girls, as well as more likely to anger, to alibi, and to give an indirect response to fear and frustration; but, even in these traits, with the exception of alibiing, sex interacts with the other independent variables.

Some of our findings we think are a function of the age group studied. This consideration and the complexity of the interrelationships uncovered suggests the need for much more exploring of the question of the effects of sibling characteristics on the personalities of children.

REFERENCES

1. CONRAD, H. S. *The California Behavior Inventory for Nursery School Children.* Berkeley: Univer. of California Press, 1933.
2. CONRAD, H. S. A statistical study of ratings on the California Behavior Inventory for Nursery School Children. *Genet. Psychol. Monogr.,* 1934, 16, 1-79.
3. EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. FOSTER, S. A study of the personality make-up and social setting of fifty jealous children. *Ment. Hyg.,* 1927, 11, 53-77.
5. GOODENOUGH, FLORENCE L., & ANDERSON, J. E. *Experimental child study.* New York: Century, 1931. Pp. 501-512.
6. KINGSLEY, ALICE, & REYNOLDS, E. L. The relation of illness patterns in childhood to ordinal position in the family. *J. Pediat.,* 1949, 35, 17-24.
7. KOCH, HELEN L. A study of some factors conditioning the social distance between the sexes. *J. soc. Psychol.,* 1944, 20, 79-107.
8. KOCH, HELEN L. The relation of "primary mental abilities" in five- and six-year-olds to sex of child and characteristics of his sibling. *Child Develpm.,* 1954, 25, 209-223.
9. KOCH, HELEN L. The relation of certain family constellation characteristics and the attitudes of children toward adults. *Child Develpm.,* 1955, 26, 13-40.
10. KOCH, HELEN L. Attitudes of young children toward their peers as related to certain characteristics of their sibling. *Psychol. Monogr.,* 1956, 70, No. 19 (Whole No. 426).
11. KOCH, HELEN L. Children's work attitudes and sibling characteristics. *Child Develpm.,* 1956, 27, 289-310.
12. LASKO, JOAN K. Parent behavior toward first and second children. *Genet. Psychol. Monogr.,* 1954, 49, 97-137.
13. MAYER, A. *Chicago: Socio-economic groups of the population by census tracts.* The Chicago Community Inventory. Chicago: Univer. of Chicago, 1940.
14. RICHARDS, T. W., & POWELL, MARJORIE. The Fels Child Behavior Scales. *Genet. Psychol. Monogr.,* 1941, 24, 259-311.
15. SANFORD, R. N., *et al.* Physique, personality, and scholarship. *Monogr. Soc. Res. Child Develpm.,* 1943, 8, No. 1.
16. SEARS, PAULINE S. Doll play aggression in normal young children. *Psychol. Monogr.,* 1951, 65, No. 6 (Whole No. 323).

17. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, PAULINE S. Some child rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-236.

18. SEWALL, MABEL. Two studies in sibling rivalry. I. Some causes of jealousy in young children. *Smith Coll. Stud. Soc. Work*, 1930, 1, 6-22.

19. STENDLER, CELIA B. Possible causes of overdependency in young children. *Child Develpm.*, 1954, 25, 125-147.

A METHOD FOR DETERMINING THE POLARITY OF BEHAVIOR ITEMS [1]

JOHN B. CARROLL
Harvard University
and HARRY LEVIN[2]
Cornell University

Personality theories are replete with bipolar concepts: love-hate, introversion-extraversion, masculinity-femininity, manic-depressive, pleasure-pain, and so forth. Research workers have sought to identify behavior items which can be considered indicative of one or the other pole of such bipolar traits.

The diagnostic significance of behavior items is often investigated by comparing observations made on contrasting reference groups. An example of this procedure is seen in Terman and Miles's (8) establishment of a masculinity-femininity scale by comparing the answers of men and women to a series of questionnaire items. It will often occur, however, that no clear reference groups are available to tie down the ends of the continuum. Also, the significance of a particular behavior item may be arguable, some theorists urging that it belongs to one end of the continuum, and others maintaining precisely the opposite. There is a place, therefore, for a technique which will provide some objective way of determining the scoring of behavior items when no external criterion is available.

The method presented here is applicable in the case where a sizable number (say, at least 10) of behavior items are postulated to have in com-

---

[1] This article utilizes data collected in a study entitled "Identification in Young Children," conducted by the Harvard Laboratory of Human Development, and supported by a grant from the National Institute of Mental Health (M-461).

[2] We wish to thank Gertrude Tourtellot and Lucile Bennis for their help with the computations.

CHILD DEVELOPMENT, Vol. 27, No. 4 (December, 1956)

mon some relevance to a certain bipolar trait, and where it is desirable to allocate each item to a position on this continuum. In simple terms, the method is designed to tell whether a certain item is to be scored "plus" or "minus." The source data enter into the analysis in a dichotomized form, either as the simple presence or absence of the behavior in question or as a result of cutting the response distribution at some point, preferably near the median. The chief outcome of the method is a specification of the polarity of each item, together with an indication of the extent to which each item is relevant to the underlying trait. If desired, the results may be used to select those items which appear to measure the trait most strongly, as shown by factor loadings. It is required, of course, that data on all behavior items to be considered be available on each member of a suitably large sample of individuals.

It cannot be too strongly emphasized that the method depends on the assumption that there is a single common dimension (either "pure" or complex) underlying all or most of the items, and that the total pool of items used in the computations provides the best basis, at least initially, for establishing that dimension. There are several obvious variations or extensions of the method, however, which may serve to minimize the dangers attending the assumption of a single dimension. For example, one could apply the method to an initial set of items in which one has high confidence, later adding other items on the basis of whether they correlate significantly with the scores established on the initial set. If large numbers of items are available, one could also apply the method iteratively, successively refining the sets of items which are at the extremes of the polarity or even applying the method to the residual set of items not found at the extremes.

Rationale and Procedure

Suppose we have a set of *n* variables each of which has been scored dichotomously; that is, each item can be scored as either "plus" or "minus," or as "zero" or "one," for example. It is assumed that all these *n* variables measure, with some error, a single polar trait; it is not known, however, to which ends of the trait the "pluses" and "minuses" should eventually be assigned. The "pluses" and "minuses" used in collecting the data may be solely arbitrary.

A criterion for assigning the "pluses" and "minuses" to the respective ends of the continuum is that the variance of the total scores over all variables must be maximized. The most rigorous approach to this problem would be the use of the theory of principal axes; as Horst has shown (5), the first principal axis yields the set of weights which maximizes the variance of the sum of scores. The weights can be either positive or negative, and can vary continuously. The principal axis solution, however, is generally laborious and costly, and would rarely be worthwhile for the type of problem with which we are concerned here. The present method uses, in effect, the centroid axis, which is usually reasonably close to the principal axis. (Edgerton and Kolbe's claims to the contrary [2] may be discounted because their centroid solution did not allow for sign-reflection.)

Fortunately, the basis for the method can be looked at quite simply as an application of a formula for the variance of a sum, which is as follows:

$$\sigma^2_{(x_1+x_2+\cdots+x_n)} = \sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i\sigma_j r_{ij}\,; \quad (i,j = 1, 2, \cdots, n).$$

That is, the variance of the sum of all variables (with unit weights) is the sum of all entries in the variance-covariance matrix of the variables. This variance can be altered, however, by changing the signs of some of the variables. (If *all* variables are reflected at the same time, the variance of the sum remains the same.) To maximize the total variance, therefore, it is necessary to find the best way of assigning signs to the variables. In the types of situations with which we are concerned here, it will often be found that many of the array totals of the variance-covariance matrix will be negative, thus counterbalancing some of the positive array sums and yielding a smaller total variance than might be obtained by reflection of the variables. Indeed, if all array sums were positive at the outset, it could mean that the investigator was remarkably sagacious in assigning the initial arbitrary scoring weights. Nevertheless, even when all array sums are positive, it is sometimes possible to obtain an even higher total variance by judicious reflection of pairs of variables, as will be shown below.

The recommended technique for obtaining a maximum total variance is that of sign-reflection, commonly employed in the centroid method of factor analysis. This technique, which is usually applied to matrices of factor residuals but is also applicable to initial correlation or covariance matrices, is presented by several writers on factor analysis, including Thurstone (9, pp. 165-166), Cattell (1, pp. 162-165), and Guilford (3, p. 497). Holley (4) has presented a supplementary procedure which should always be tried after sign-reflection has arrived at the stage where all array sums in the matrix are positive.

In the present case, one does not need to reduce the data to a correlation matrix, or even to a variance-covariance matrix; one can just as well operate on a matrix whose entries are the variances and covariances multiplied by N, for such a matrix is easily computed from frequencies of joint occurrence. If the entry in the ith row and the jth column of this matrix is denoted by b_{ij}, such an entry is computed by the formula

$$b_{ij} = N_{ij} - p_i N_j\,, \qquad \text{(1)}$$

where N_{ij} = the number of cases getting a "plus" or "1" in both variables i and j, p_i = the proportion of cases getting a "plus" or "1" in variable i, and N_j = the number of cases getting a "plus" or "1" in variable j.

One then applies the technique of sign-reflection to this matrix until the grand sum of all arrays is as large as possible. It is always possible to reflect until every array has a positive sum, and Holley's technique should be tried after this stage is reached in order to see whether further sign-reflections are feasible. The variables whose signs are negative at the end of this procedure are, then, the ones to be reversed in scoring.

The final step is not statistical, but solely interpretative: to examine all behaviors which are similarly scored. That is, one examines the behaviors indicated to be at each end of the continuum, in order to decide on labels for these ends. For example, if one is working on masculinity-femininity behavior items, one would examine all behaviors at one end of the continuum to see whether they are typically masculine or typically feminine. In general, items with the highest array sums are those which can be expected to give the clearest evidence on the identification of poles. One can easily, in fact, compute the centroid factor loadings in order to rank the variables with respect to their position on the underlying trait dimension, or in order to select subtests of variables having factor loadings greater, in absolute magnitude, than some chosen value.

The results of this procedure can be regarded as very clear or as only suggestive, depending on the relative size of the array sums and upon the plausibility of the interpretations. It may happen that the initial covariances are very small on the average, in which case the results may actually be due to chance rather than to any common trait measured by the behavior items. At any stage of the sign-reflection process, but most pertinently at the final stage, one may compute the reliability of the sum of variables and test its significance.

It should be pointed out that the polarity of the final scale as a whole is arbitrary; it is conceivable that in two random samples of data the positive end of the continuum in one sample would accidentally correspond to the negative end of the continuum in the other. (To understand this, it may be noted that simultaneous reflection of all variables does not alter the variance of a sum.) Therefore, in cases where one is comparing or contrasting two or more samples it may be desirable to reflect all signs in certain samples; such reflection is further necessary, if one is establishing a scoring system, to insure that the scores will have comparable meanings in all samples.

COMPUTATIONS AND FORMULAS

The procedures and computations will now be illustrated in detail for a small example with six variables, as shown in Table 1. (These data are extracted from the data of a larger study whose results are presented in Table 2; in practice, it is advisable to have at least 10 variables in the analysis). The initial matrix B consists of entries b_{ij} as defined in formula (1).

TABLE I

ILLUSTRATIVE COMPUTATIONS ($N = 202$)

			(—)			(—)	
	9	12	14	17	33	37	Σ
9	(50.4)	2.3	1.9	—1.6	23.1	—6.8	18.9
12	2.3	(48.1)	—11.8	4.2	3.9	—7.2	—8.6
14	1.9	—11.8	(50.3)	—7.6	3.6	0.0	—13.9
17	—1.6	4.2	—7.6	(46.1)	—1.4	—7.8	—14.2
33	23.1	3.9	3.6	—1.4	(50.4)	—16.0	13.2
37	—6.8	—7.2	0.0	—7.8	—16.0	(49.2)	—37.8
							$\Sigma(\text{diag.}) = 294.5 = D$
S	18.9	—8.6	—13.9	—14.2	13.2	—37.8	—42.4
$-\frac{1}{2}S$	—9.45	4.30	6.95	7.10	—6.60	18.90	21.20
+37	—16.25	—2.90	6.95	—0.70	—22.60	*18.90*	—16.60
+14	—14.35	—14.70	*6.95*	—8.30	—19.00	*18.90*	—30.50
C	28.7	29.4	*13.9*	16.6	38.0	*37.8*	$164.40 = A$
T	79.1	77.5	*64.2*	62.7	88.4	*87.0*	$458.9 = A + D$
a	.260	.254	—.211	.206	.290	—.286	$\sqrt{(202)(458.9)} = 304.5$

$$r_t = \frac{6}{5} \cdot \frac{164.4}{294.5 + 164.4} = .430$$

$$F = \frac{164.4}{294.5 - (258.9)/6} = .75 \quad ; \quad p > .50$$

$$d.f._1 = 202 - 1 = 201$$

$$d.f._2 = (202 - 1)(6 - 1) = 1005$$

It is well to indicate the diagonal entries in some special way (here they are put in parentheses), as these are not used in the sign-reflection process. The computation steps are:

1. Set up a check column to the right of the matrix B, entries being the sums of nondiagonal entries for each row in the matrix.
2. Add the nondiagonal entries for each column and record in the row labeled S.
3. Divide each entry in S by 2 and reverse its sign, recording the results in the row labeled $-\frac{1}{2}S$.

4. Identify the variable having the highest positive entry in the row labeled $-\frac{1}{2}S$; place a negative sign at the head of its column, then find that variable's *row* in the matrix and add its entries to those of the row labeled row $-\frac{1}{2}S$, recording the results in the row immediately below. Add across and check the result in the check column. The diagonal entry in the matrix is to be regarded as zero, so that in the row below $-\frac{1}{2}S$ the entry corresponding to the diagonal position is the same as the entry in row $(-\frac{1}{2}S)$. (In Table 1, the results are recorded in row "+37," since variable 37 was reflected.) At the same time mark this entry in some special way, e.g., with red pencil, and carry this indication down as long as this variable remains reflected. (Here the special marking is indicated by the use of *italics*.)

5. In the row just recorded, identify the variable having the highest positive total, place a negative sign at the head of its column, then find the variable's row in the matrix and add its entries to those of the row just recorded. The results are recorded in a new row (row "+14" in the example).

6. Continue in this way until no *positive* entries remain for variables *not* reflected, and no *negative* entries for variables *reflected*. (In the case of a negative entry for a reflected variable, that variable's row is subtracted instead of added, and the variable is subsequently treated as an unreflected variable.)

7. Inspect the computations for the possibility that further reflections can be made by reflecting *pairs* of variables (4). This will be the case if an entry in the ith row and the jth column, after it itself has been changed in sign in accordance with any reflections made up to this time, is greater than the absolute sum of the entries in the ith and jth columns of the last row (assuming that that row conforms to the requirements of step 7.) For example, in Table 1 if $b_{(9)(33)}$, which is 23.1, were larger than $|14.35 + 19.00|$, the absolute sum of the corresponding entries in row "+14," one would proceed to reflect first variable 9, then variable 33. However, since this is not the case, this step cannot be illustrated in this example.

8. Being satisfied that no further reflections are possible, multiply each entry in the last increment row by 2 and reverse the signs of the entries for variables not reflected, recording the results in the row labeled C. Row C's entries should thus all be positive. The object of the sign-reflection procedure is to maximize the sum of these entries, which are the columnar sums of the entries in the computation matrix B after sign-reflection. Add across row C to find the sum of the entries; this sum is designated A and is used in later work.

9. Find the sum of the diagonal entries of matrix B, heretofore not used. This sum is D.

10. Add the diagonal entries of B to the entries of row C, column by column, to form row T. Let t_j be the jth entry of row T. Then $\Sigma t_j = (A + D)$, and is equal to N times the variance of the total scores on all n variables when they are added together with weights of $+1$ for unreflected variables and -1 for reflected variables (or the exact opposite).

11. The reliability of the total score obtained by adding together the scores of "1" or "0" on the variables, with unit weights and signs determined by the preceding analysis, may now be computed by the formula

$$r_t = \frac{n}{n-1}\left(\frac{A}{A+D}\right) \text{ . } \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

To test the significance of this reliability under the hypothesis of zero reliability in the population, evaluate the following F ratio:

$$F = \frac{A}{D-(A+D)\,/\,n} \text{ , } \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

with $d.f._1 = (N-1)$ and $d.f._2 = (N-1)(n-1)$. (These formulae are derived and adapted from those presented by Hoyt [6] for the case of multiple-item tests.)

12. If desired, compute the final centroid loading for any variable j by the formula

$$a_j = \frac{t_j}{\sqrt{N\Sigma t_j}} \text{ . } \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

(This uses unities in the diagonal of the correlation matrix corresponding to matrix B. Alternatively, one may replace the diagonal elements of matrix B with values which will be $N\sigma^2{}_j$ times the estimated communalities of the variables. For example, one may use the highest value in each array as a substitute for the diagonal entry.)

If a sufficient number of cases are available, it may be advisable to apply the procedure to two random halves of the sample in order to note whether the results are consistent.

ILLUSTRATIVE EXAMPLE

The need for the technique outlined above became evident in the course of studies now under way at the Harvard Laboratory of Human Development. These studies are concerned with the child-rearing antecedents of identification in young children. One aspect of identification is the extent of super-ego development. The function of the super-ego, according to psychoanalytic theory, is to control the impulsive behavior of the individual. It is an end-product of the child's identification with his parents, and represents an internalization of the parents' values as to what are right and wrong behaviors. In addition, a child with a well-developed super-ego has taken over from his parents the prerogative of punishing himself for wrongdoing as well as that of rewarding himself for behaving in a manner which the parents heretofore approved and rewarded. In other words, one significant aspect of super-ego development is a shift in the control of behavior from control by rewards and punishments administered by some agent in the child's environment to internal controls which operate even when the external controlling situation does not operate or ceases to exist.

With this theory as a background, super-ego development is thought to be amenable to operational definition and measurement in terms of the number and kind of responses the child makes to a series of childhood "deviations" or wrongdoings presented to him in a doll-play, make-believe situation.[3] It was postulated that children's responses in this situation could be ordered along a continuum ranging from "guilt" to "fear," the former being represented by instances of punishment to the self, and the latter by such signs as escape from a potential, external punisher. To the extent that the child gives responses manifesting "guilt," the child can be said to exhibit super-ego development. It is also conceivable that super-ego development would be represented by instances in which the child depicts adult-like behaviors, and thus, we may infer, shows identification with parents.

The responses to the doll-play situation could, of course, be scored and interpreted purely in an *a priori* manner on the basis of the theory outlined above. Such a procedure would be highly subjective, however, and is hardly to be preferred over any method which would appeal to the actual nature of the data. The method described in the first part of this paper was developed for the purpose of establishing a bipolar continuum and assigning the behavior responses to one end or the other of this continuum. It was necessary to assume only that all the items could be described in terms of *some* continuum, but it was not necessarily assumed that this continuum was a guilt-fear continuum.

The data used in the present study were behavioral observations of children collected in a situation designed specifically to elicit acts which could be ordered theoretically along a bipolar dimension which was tentatively called "guilt-fear." The subjects were 202 boys and 177 girls, all five years old. Prior to the guilt-fear observation session, all the children had been observed in two sessions of regular doll play, by a method which is described elsewhere (7). For the guilt-fear session, the experimenter took the child from the kindergarten into an experimental room where there was set up an uncovered, furnished six-room dollhouse, in which a family (father, mother, boy, girl, baby) of small flexible dolls could be manipulated. The experimenter then acted out for the child a series of eight deviations, or childhood wrongdoings, using as the "deviator" or wrongdoer a child doll of the same sex as the subject. An example of one of the eight deviation situations follows:

> The mother puts the baby on the couch. The mother tells the boy to watch the baby while she goes out. The mother tells the boy to watch that the baby does not fall off the couch. The mother goes out. The boy stays for a while. The boy goes out and plays. The baby falls off the couch. What happens now?

---

[3] This doll play technique for investigating guilt and fear was developed by Drs. Pauline S. Sears (now of Stanford University) and Edgar L. Lowell (now of the John Tracy Clinic, Los Angeles), both formerly of the Harvard Laboratory of Human Development. A complete report is in preparation.

The subject was then permitted to complete the story using any members of the doll family that he chose. The experimenter recorded as many acts as the child used to finish the story. The subject's responses were recorded in a prearranged set of categories, as follows: (1) "fixing" (e.g., the boy put the baby back on the couch); (2) physical punishment (the mother spanked the boy); (3) verbal punishment (the mother says, "You've been a naughty boy"); (4) other punishment (the mother made the boy go to bed); (5) hurt feelings (the boy felt sad); (6) delay (the subject delays giving any relevant response); (7) redefinition (the baby really didn't fall down);

TABLE 2

FACTOR LOADINGS OF BEHAVIOR ITEMS FOR TOTAL SAMPLE AND RANDOM HALF-SAMPLES

	Boys (N=202)			*Girls* (N=177)		
Variable	Total	I	II	Total	I	II
37 Total Unspecified Agent	.25	.27	.30	.22	.17	.22
36 Total Girl Agent	.19	.23	.14	.27	.27	.32
29 Cs[*] Agent of "Other" Punishment	.19	.12	.22	.20	.13	.24
20 "Hiding"	.18	.16	.20	.12	.16	.09
8 Cs[*] Punished by Unspecified Agent	.18	.18	.19	.23	.19	.22
6 Physical Punishment, Cs[*] to Mother	.17	.16	.17	.09	.06	.10
16 "Feeling Hurt"	.16	.20	.11	.10	—.09[†]	.10[†]
14 Verbal Punishment	.15	.16	.15	—.15	—.23	—.11
10 Physical Punishment, Father to Cs[*]	.14	.14	.21	—.17	—.16	—.20
19 "Confession"	.13	.17[†]	—.11[†]	—.14	—.13	—.15
18 "Redefinition"	.13	.19[†]	—.11[†]	.17	.09	.23
35 Total Boy Agent	.10	.10	.10	.08	—.12[†]	.13[†]
34 Total Father Agent	.03	—.02[†]	.06[†]	—.16	—.17	—.16
17 "Delay"	—.17	—.17	—.19	.17	.21	.15
9 Physical Punishment, Mother to Cs[*]	—.18	—.17	—.14	—.22	—.27	—.22
12 "Fixing"	—.18	—.27	—.20	—.10	.11[†]	—.11[†]
28 Mother Agent, "Other" Punishment	—.18	—.11	—.19	—.29	—.29	—.30
33 Total Mother Agent	—.31	—.23	—.28	—.28	—.29	—.25
r_t	.58	.59	.60	.61	.61	.64
F	1.38	1.41	1.48	1.58	1.57	1.81
$d.f._1$	201	100	100	176	88	87
$d.f._2$	3417	1700	1700	3009	1496	1479
p	$<.01$	$<.01$	$<.01$	$<.01$	$<.01$	$<.01$

[*] Child doll of same sex as subject.

[†] Discrepancy between samples in sign of factor loading.

(8) confession (the boy told the mother what he has done); and (9) hiding (the boy ran away).

From these data, a large number of variables could be derived, depending on the type of act, the agent and object of the act, etc. The 18 variables listed in Table 2 were selected for study as being theoretically most relevant to the guilt-fear dimension, if it existed. In preparing the data for analysis, the frequency distributions were dichotomized as near their medians as possible; the matrix B was then computed by formula (1). Theoretical considerations and previous results from doll play with children led us to expect different patterns of responses for boys and girls. Consequently, the analysis was done separately for each sex. Also, to provide cross-validations the cases were randomly divided into two groups each for boys and for girls; there were thus six separate analyses in all.

The results of applying the bipolar scoring method are presented in Table 2. The variables are rank-ordered in the table from highest positive to highest negative factor loadings on the basis of the results from "Boys' Total Sample." All factor loadings for the three girls' samples were reversed in sign, since as pointed out above, the polarity of the scale as a whole is arbitrary, and it appeared that the boys' and girls' results could be made to correspond more closely by reversing the signs of the results for one of the sexes.

The reliabilities of the total scores in each of the samples are adequate, and highly significant according to the F test provided by formula (3). In general, the signs and magnitudes of the factor loadings are consistent between random samples of each sex, and even between sexes—though to a lesser extent. For each sex, there are only three variables for which the signs of the factor loadings did not agree in the two samples; these factor loadings are indicated by a dagger (†) in Table 2. In other words, in three of the 18 variables for each sample, the assignment of polarity is obviously questionable. This may be due to unreliability in the measures themselves or to a possible multi-dimensionality in this set of measures.

The final step is the interpretation of the results. This involves the inspection of the positive and negative groupings with a view to identifying what continuum appears to be measured and determining the polarity of this continuum. Variables with the higher factor loadings should be given more weight in this process.

For the boys, not all of the variables at either end of the continuum are interpretable in terms of our original theoretical formulation. The variables with the five highest negative factor loadings, however, seem to be meaningful as guilt indices. As numbered in Table 2, these variables are 33, 28, 12, 9, and 17. The strongest clue to guilt is variable 9, which is the relative amount of physical punishment which the subject represents as being given by the mother doll to the boy doll. Variable 28 again presents the mother as the agent of punishment, this time punishment other than physical and verbal; nevertheless, as in variable 9, the punishment is usually directed to the deviator. Variable 33, the proportion of scoring units that depict the mother doll as the agent of action, is interpretable as an index of the extent to which the subject identifies with the mother. Variables 17 and 12 are most easily interpeted as indices of guilt: the "fixing" response of variable 12 represents an acceptance by the child that there has been a wrongdoing and he attempts to make retribution by undoing the wrong; "delay" (variable 17) indicates that the child, overwhelmed by the deviation, manifests a "block" so strong that he makes no response whatever, at least initially. In summary, four of the five variables with high negative factor loadings are interpretable as guilt indices and the fifth appears to be an index of identification. This finding supports the theory that guilt is a characteristic of a well-developed super-ego, which in turn is associated with parent-identification.

Variables with high positive factor loadings for the boys are those numbered 37, 36, 29, 20, 8, and 6. "Hiding" (variable 20) is possibly a a sign of fear, in that the deviator is represented as trying to avoid punishment for the transgression. Variables 29 and 6 represent counter-aggression in response to the deviation; that is, the boys add further transgressions to those already depicted. Variables 37 and 36 are difficult to interpret; as indices of actions by agents (either the girl doll or "unspecified" agents) other than the boy doll, they do not at any rate appear to measure behaviors which we would expect guilty children to show, nor do they appear to measure parent-identification. Variable 8 is particularly difficult to interpret as any kind of fear index, since it represents punishment directed against the deviator by "unspecified agents." On the whole, the positively signed variables do not seem to represent the "fear" end of a guilt-fear continuum as originally postulated; rather, they seem to constitute merely the opposites of the guilt and identification behaviors isolated at the other end of the continuum. Possibly the continuum may be thought of as ranging from "identification and guilt" (at the negative end) to "nonguilt" (at the positive end), or perhaps the continuum represents two mutually exclusive ways of evidencing guilt.

The findings for the girls present a generally similar appearance. As in the case of the boys' results, the negative end of the continuum is clearer. The five measures with the highest negative factor loadings, those numbered 28, 33, 9, 10, and 34, seem to represent guilt, in that the action is portrayed through a parent with whom the child is identifying or the action represents the child deviator being punished. The positive end of the continuum, however, is difficult to interpret as representing fear. Some of the variables with positive factor loadings, such as variables 29, 37, and 36, seem to indicate absence of identification with parents. Only "redefinition," variable 18, is an expected sign of fear. It is notable that variable 17, "delay," which was at the negative end of the continuum for boys, grouped with the positively-loaded variables for the girls. The results for the girls seem to confirm the conclusion reached for the boys' data, namely, that the negative end of the continuum represents guilt and parent-identification, whereas the positive end indicates the absence of these behaviors, or possibly even the rejection of guilt feelings.

It should be pointed out that the data were collected as a part of a larger study designed to relate certain child-training measures to evidences of guilt and fear in children. The method of analysis described here provides a first approximation of guilt and identification scores. It is hoped that the relationships will be further clarified when the data from the child-training measures are analyzed.

REFERENCES

1. CATTELL, R. B. *Factor analysis; an introduction and manual for the psychologist and social scientist.* New York: Harper, 1952.
2. EDGERTON, H. A., & KOLBE, L. E. The method of minimum variation for the combination of criteria. *Psychometrika,* 1936, 1, 183-187.
3. GUILFORD, J. P. *Psychometric methods.* (2nd Ed.) New York: McGraw-Hill, 1954.
4. HOLLEY, J. W. A note on the reflection of signs in the extraction of centroid factors. *Psychometrika,* 1947, 12, 263-265.
5. HORST, A. P. Obtaining a composite measure from different measures of the same attribute. *Psychometrika,* 1936, 1, 53-60.
6. HOYT, CYRIL. Test reliability obtained by analysis of variance. *Psychometrika,* 1941, 6, 153-160.
7. SEARS, PAULINE S. Doll play aggression in normal young children: influence of sex, age, sibling status, father's absence. *Psychol. Monogr.,* 1951, 65(6).
8. TERMAN, L. M., & MILES, C. C. *Sex and personality.* New York: McGraw-Hill, 1936.
9. THURSTONE, L. L. *Multiple factor analysis.* Chicago: Univer. of Chicago Press, 1947.

THE EFFECT OF NONDIFFERENTIAL REWARD AND NON-REWARD ON DISCRIMINATIVE LEARNING IN CHILDREN[1]

ALLEN D. CALVIN, M. J. PERKINS *and* FREDRIC K. HOFFMANN
Michigan State University

In a recent paper Calvin (5) pointed out that his findings on the effect of nondifferential reward and punishment could not be accounted for by Hull's (6) reinforcement theory or Bitterman's (2) concept of perceptual organization, and that some revision of both Hull's and Bitterman's positions seemed necessary. Since that time Bitterman (3) has revised his position considerably to enable him to handle adequately recent experimental findings. The present experiments were designed to extend the investigation of the effect of nondifferential experience utilizing children for the experimental population.

EXPERIMENT I

Apparatus and Procedure

A 28 by 22 in. cardboard screen (which could be raised and lowered for each trial, and which was set in a masking framework of cardboard extending a foot on either side) was used to present all the training and testing trials to the *S*s who were seated beyond arms length on the opposite side of the screen from *E*. Two identical plain opaque red plastic cups were presented mouth down and 4 in. apart. If *S* pointed to the cup with the reward (a small plastic toy of the kind given in gum ball machines) under it, *E* picked up the cup and gave the reward to *S*; whereas if *S* pointed to the cup with nothing under it, *E* first picked up the chosen cup to show that it had not covered a reward and then immediately lowered

---

[1] The authors wish to express their thanks to C. E. MacDonald, Superintendent of Schools, East Lansing, Michigan; Noel M. Ranger, Principal of the Red Cedar School in East Lansing; and Brewster Lewis, Principal of the Bailey Street School in East Lansing; for their kind cooperation which made these studies possible.

The data in the first experiment were presented at the 1954 APA convention in New York. The second experiment was supported by a grant from the Social Science Research Council.

The senior author is now at Hollins College.

CHILD DEVELOPMENT, Vol. 27, No. 4 (December, 1956)

the screen to terminate the trial without giving *S* a reward. Two 3 by 4½ in. white cardboard stimulus cards which served as cues were placed in horizontal positions, one directly in front of each cup.

The *S*s were 36 elementary public school children who were divided into three groups consisting of 12 *S*s. Half of each group was randomly chosen from the fifth grade and half from the sixth, with the only restriction being that they be equally made up of males and females. Group I (the control group) had to solve a standard discrimination problem. They had to learn to choose the correct stimulus regardless of the spatial location, which was varied according to a prearranged random order such that neither stimulus appeared more than three times in succession in either position. One stimulus card had three ½ in. medium metallic green gummed stars spaced at ⅝ in. intervals along the center line of the length of the card, whereas the other card had four such stars along the center line. The three starred card was positive for half the *S*s and the four starred positive for the other half. Group II (the nondifferential reward group) was given 20 training trials in which a plain white card was always negative and a starred card always positive; i.e., on half the trials *S* found white vs. three stars, and on the other half, white vs. four stars, but the starred card was always positive. Upon completion of the training trials, *S*'s from this group were switched to the test problem that Group I worked on, three vs. four stars. Group III (the nondifferential non-reward group) was given 20 training trials in which the starred cards were negative and a plain white card was positive, the reverse of Group II. At the conclusion of the training trials this group also was switched to Group I's problem. A balanced experimental design was used in the presentation of the stimuli to all groups.

The criterion of learning on the test problem was 10 consecutive correct trials. The test problem was terminated at the tenth consecutive correct trial or at the first error after the thirtieth trial, whichever occurred first.

Results and Discussion

Table 1 presents median trials to learn for all groups. Using the Mann-Whitney U Test (8), the following results were obtained (all tests were two-tailed): the difference between Group I, the control group, and Group III, the nondifferential non-reward group, was significant beyond the 5 per cent level; the difference between Group I and Group II, the nondifferential reward group, was significant beyond the 7 per cent level; the difference between Group II and Group III fell far short of significance. Thus the groups that were given nondifferential experience were superior to the control group, and there was no appreciable difference between the nondifferential reward group and the nondifferential non-reward group.

When we compare these findings with those of other experimenters who used rats as *S*s, we note quite a bit of similarity and some differences. The superiority of our nondifferential reward group to the control group is in general agreement with the findings of Bitterman, Calvin, and Elam (2) and Bitterman and Elam (3). Billingsley, Feddersen, and Bitterman (1), Bitterman, Elam, and Wortz (4), Calvin (5), and Siegel (9) did not find this superiority. In an experiment in this area using adult humans as *Ss*, Maltzman and Morrissett (7) also failed to find results corresponding to ours. A possible reason for these discrepant findings will be discussed below.

TABLE I

MEDIAN TRIALS TO LEARN

Group	*Trials*	*No. of Ss Who Failed to Solve*
I. Controls	20	5
II. Nondifferential reward	13	2
III. Nondifferential non-reward	11	3

The fact that the performance of our nondifferential reward and nondifferential non-reward groups are not significantly different is in accord with the findings of Bitterman, Elam and Wortz (4), but Calvin (5) found that his nondifferential reward group was significantly superior to his nondifferential punishment group. The superiority of the nondifferential non-reward group to the control is not in accord with the findings of either Bitterman, Elam, and Wortz or Calvin. However, as Calvin pointed out, the findings he obtained indicated that punishment of the kind used in jumping stand experiments introduces an extraneous factor such as frustration or rigidity which makes evaluation of the actual effect of nondifferential non-reward impossible in this type of experimental situation. In our experiment it was possible to avoid the introduction of such a factor, by using non-reward rather than punishment, and it is this modification in experimental design rather than any phylogenetic change which the present authors feel accounts for the difference between our findings and those of other investigators who used rats for their experimental population.

Before attempting to resolve the discrepancy between those findings which indicate superior performance for groups having nondifferential experience and those experiments which fail to confirm this hypothesis, let us examine for a moment the theoretical positions of Hull and Bitterman which prompted much of the research in this area. Hull's (6) concept of a negatively accelerated growth function for habit strength, of course, leads to the prediction that it is more difficult to establish a differential response between two strong habit patterns than between two weak ones. [For a more detailed derivation see Hull (6, p. 29).] Bitterman has recently modified his position. He has discarded his original perceptual differentiation hypothesis (2) for a two factor explanation (3) which may be summarized as follows: (*a*) a general tendency for nondifferential reward to retard future differential discrimination due probably to some attentional modification; (*b*) nondifferential reward on the actual stimuli to be discriminated lessens the general tendency for nondifferential reward to retard —either by perceptual differentiation or some other factor which serves to forestall the negative transfer.

At first glance, both Hull's and Bitterman's theoretical positions seem unable to account for our results. Since Bitterman has posited negative transfer from a nondifferential reward situation to a discrimination situation (although this negative transfer is ameliorated somewhat by exposure to the same stimuli to be discriminated), he would not be able to account for the inferior performance of the control *S*s in Group I who had *no* nondifferential experience. It appears that since Hull has postulated a negatively accelerated growth of habit strength, he too would be unable to explain the inferior performance of Group I. However, on closer inspection, we note that during their predifferential training Groups II and III were working on a problem similar to the test problem. Thus positive transfer from the training problem to the test problem may well account for the superior performance of these *S*s. That some such factor is operating receives considerable support when we examine the effect which performance on the preliminary problem has on performance on the test problem. Of the seven *S*s who failed to solve the preliminary problem, four failed to solve the test problem. A breakdown of the 17 *S*s who solved the preliminary problem revealed that only two failed to solve the test problem. (An *S* was considered to have solved the preliminary problem if his last eight choices were correct.) A χ^2 was computed, and, using Yates' correction for continuity, a χ^2 of 4.82 was obtained which is significant beyond the 5 per cent level. When the actual scores on the test problem of those *S*s who failed the preliminary problem are compared by the U test with the test problem scores of those *S*s who solved the preliminary problem, the difference is significant beyond the 1 per cent level. Thus, those *S*s who solved the preliminary problem were significantly superior on the test problem. Since approximately 70 per cent of the *S*s in Groups II and III solved the preliminary problem, it seems reasonable to assume that the superiority of these groups on the test problem was due in a large degree to positive transfer from the preliminary problem.

When viewed in the light of the preceding analysis, it seemed to the authors that it might prove fruitful to analyze previous experiments in nondifferential reward and punishment in terms of negative and positive transfer from the preliminary problem to the test problem. After completing such an analysis, it appears to us that the discrepant findings which seemed to necessitate a number of *ad hoc* assumptions for explanation could be accounted for by the following postulates relating to transfer: (*a*) In a situation where two stimulus objects receive nondifferential reward or punishment[2] in a context where there is an available solution, this training will result in positive transfer. The amount of positive transfer will be inversely related to the number of possible solutions available in the nondifferential situation (assuming, of course, that the problem has only one actual solution) and will be directly related to the degree to which these possible solutions have been explored. (*b*) In a situation where two stimulus objects receive nondifferential reward or punishment in a context where there is no available solution, this training will result in negative transfer. The amount of negative transfer will be inversely related to the number of possible solutions available in the nondifferential situation (in this case, of course, none of the "solutions" will actually solve the problem) and will be directly related to the degree to which these alternative solutions have been explored.

To simplify matters, we will assume that in the course of nondifferential experience each possible solution has an equal chance of occurring. When this is not the case, of course, then the number of possible solutions cannot be merely added together, but instead each solution must be weighted according to its relative probability of occurrence before they are combined. A situation where the probability of occurrence of the actual solution is very low in comparison to the other alternative solutions will necessitate a great deal of nondifferential experience before significant positive transfer to the new situation will take place, etc.

It also appears likely that the amount of transfer would not prove to be a simple linear function and that it would approach an asymptote relatively unaffected by either the degree of exploration or the number of possible solutions; however, the exact shape of the curve will have to await additional experimental work.

Space will not permit us in the present paper to present our analysis of the previous work in this area utilizing empirical findings to support our theoretical formulation, but it seems apparent that without controlling for the effects of transfer from the nondifferential situation to the test problem an adequate test of such concepts as the growth of $_{S}H_{R}$, perceptual differentiation, etc., is not possible.

EXPERIMENT II

This experiment was designed to control for the factor of transfer in nondifferential experience; thereby enabling us to obtain a test of Hull's and Bitterman's positions which would not be confounded as in previous experiments which did not control for this factor.

---

[2] Punishment here includes both non-rewarded situations and situations in which the *S* is actually punished for making an incorrect response. In the latter case, as Calvin (5) pointed out, it is often impossible to evaluate adequately the effect of nondifferential experience.

Method

The *S*s were 48 children from the third grade of a local public school. They were randomly divided into three groups of 16 each, the only restriction being that each group consist of eight males and eight females.

The apparatus and procedure were similar to that utilized in Experiment I.

Group I (the Control Group) was presented with a standard simultaneous discrimination problem with the three starred and four starred cards described in Experiment I serving as cues. The three starred card was positive for half the *S*s and the four starred positive for the other half. Group II (the Same Group) received 20 training trials of white vs. three and four stars. The appearance of the three or four starred card on a given trial was determined according to a modified random order so that each card was presented no more than three times in succession, and each was viewed equally often. For one-half the *S*s the starred cards were positive and for the other half the white was positive. Upon completion of the training trials *S*s were given the standard discrimination problem of Group I, i.e., to discriminate between three vs. four stars. Group III (the Different Group) received 20 training trials in the same manner as Group II except that instead of the starred cards they were shown two figured cards—one with a green triangle in the center and the other a green circle. As in the case of Group II the appearance of the triangle or circle card on a given trial was determined according to a modified random order so that each card was presented no more than three times in succession, and each was viewed equally often. For one-half the *S*s the figured cards were positive and for the other half the white card was positive. Upon completion of the training trials *S*s in Group III were given the standard discrimination problem of Group I. The criterion for learning was the same as in Experiment I.

Results and Discussion

Since the *S*s in Group II and III had training on a problem which should lead to equal amounts of transfer in terms of our previous postulates, let us see what Bitterman and Hull would predict in regard to our findings. Bitterman's position, as we noted previously, maintains that there is a general tendency for nondifferential reward to retard future differential discrimination due probably to some attentional modification, and that nondifferential reward on the actual stimuli to be discriminated lessens the general tendency for nondifferential reward to retard—either by perceptual differentiation or some other factor which serves to forestall the negative transfer. Thus he would predict that Group I should perform better than either Groups II or III and further that Group II's performance would prove superior to Group III.[3] Hull does not deal specifically with comparisons of Group I contrasted with Groups II and III, but he does predict that the performance of Group III should be better than that of Group II. Since there was no difference in performance between those *Ss* receiving nondifferential reward and those receiving nondifferential non-reward, these two groups were combined and the general findings are presented in Table 2. The trials to learn were tested by the U test (all tests were two-tailed). The difference between Group I and Group III was significant beyond the 5 per cent level; the difference between Group I and Group II was significant beyond the 8 per cent level; and the difference between Group II and Group III was not significant. Since none of the predictions made from Bitterman's position were confirmed, it would seem that some radical change in his theory is called for. The performance of Group III was superior to Group II as predicted from Hull's theoretical position; however, this difference was not significant. The authors believe (although this is admittedly *ad hoc*) that this lack of significance can be ascribed to the relative position on the generalization gradient of the test and training stimuli; therefore, we are about to begin an experiment in which the stimuli used are much further apart on the generalization gradient, thereby lessening the effect of stimulus generalization.

TABLE 2

MEDIAN TRIALS TO LEARN

Group	*Trials*	*No. of Ss Who Failed to Solve*
I. Controls	13	5
II. Same	11	3
III. Different	10	0

We believe our findings indicate the necessity of taking the type of preliminary training into account when designing experiments to test such concepts as perceptual differentiation, $_SH_R$, etc. In previous experiments presenting results favorable to Bitterman's theoretical formulation (2, 3), the control animals were presented with two identical grey cards thus cutting down the amount of possible solutions during the training period and, according to our postulates, increasing the relative amount of negative transfer for the controls. When this variable is controlled, as in the present experiment, the results fail to support Bitterman's position.

---

[3] Bitterman does not specifically deal with the non-reward situation. However, the trends for both the rewarded and non-rewarded *Ss* were the same.

SUMMARY AND CONCLUSIONS

Two experiments are presented which indicate the importance of the type of preliminary training in evaluating the effect of nondifferential experience. When the amount of possible choices in the nondifferential training is held constant, the results obtained fail to support Bitterman's recent theoretical formulations. Hull's position also is not supported in a statistically significant manner. The implications of these findings are discussed.

REFERENCES

1. BILLINGSLEY, A. B., FEDDERSEN, E. E., & BITTERMAN, M. E. Discrimination following nondifferential reinforcement with differential afferent consequences. *Amer. J. Psychol.*, 1954, 67, 335-337.
2. BITTERMAN, M. E., CALVIN, A. D., & ELAM, C. Perceptual differentiation in the course of nondifferential reinforcement. *J. comp. physiol. Psychol.*, 1953, 46, 393-397.
3. BITTERMAN, M. E., & ELAM, C. B. Discrimination following varying amounts of nondifferential reinforcement. *Amer. J. Psychol* 1954, 67, 133-137.
4. BITTERMAN, M. E., ELAM, C. B., & WORTZ, E. C. Perceptual differentiation as a function of nondifferential reward and punishment. *J. comp. physiol. Psychol.*, 1953, 47, 475-478.
5. CALVIN, A. D. The growth of learning during nondifferential reinforcement. *J. exp. Psychol.*, 1953, 46, 248-254.
6. HULL, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
7. MALTZMAN, I., & MORRISETT, L., JR. Different strengths of set in the solution of anagrams. *J. exp. Psychol.*, 1952, 44, 242-246.
8. MANN, H. B., & WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, 18, 50-60.
9. SIEGEL, P. S. The role of absolute initial response strength in simple trial-and-error learning. *J. exp. Psychol.*, 1945, 35, 199-205.

THE DIRECTION OF AGGRESSION IN THE MOTHER-CHILD PUNISHMENT SITUATION[1]

PATRICIA K. MORGAN *and* EUGENE L. GAIER
Louisiana State University

While punishment has generally been held to be a mechanism by which socialization takes place (3, 4, 7, 9, 12, 13) and has been offered as an explanation or a contributing factor in behavior problems ranging from delinquency (8, 13) to psychosis, the majority of investigations studying this phenomenon have been confined to animal experimentation. The inability and/or lack of feasibility of placing human subjects into a punishment siutation, combined with the lack of available techniques, have doubtless served to hinder research with humans. Several major techniques have been commonly employed to investigate punishment on the human level. Through use of records of the TAT, sentence completion test responses, and attitude questionnaires, Hoffman (5) attempted to relate reactions to parental punishment to compulsive conformity. Sears and his associates (3, 11, 12) made extensive use of doll-play interview and rating scales in their investigations of frustration and punishment.

The purpose of the present study was to investigate the usefulness of a newly-designed projective device—the Punishment Situations Index (PSI)—for assessing punishment in the mother-child relationship. Specifically, we are interested in examining the perception of the direction of aggression in the punishment situation.

THE PUNISHMENT SITUATION INDEX

The Punishment Situation Index consists of two sets of 10 cartoons each, one set for use with girls and another for use with boys. In each picture, a child and his (her) mother are depicted in a situation commonly

---

[1] This research study was supported by funds made available by the Federal Research Project on Family Life. The writers wish to acknowledge the kind cooperation given by Dr. Clara Tucker, Project Director.

CHILD DEVELOPMENT, Vol. 27, No. 4 (December, 1956)

followed by punishment. Card 6 from both series is shown in Figures 1 and 2. In order of presentation, the 10 situations are as follows:

Card I. Situations involving possible physical injury. Male—Boy is shown emerging from a fight with two other boys. Female—Girl is hanging by her knees from a tree.

Card II. Unfavorable relationships with siblings. Male—Boy is teasing his sister while she cries. Female—Girl has taken her little brother's toy away from him.

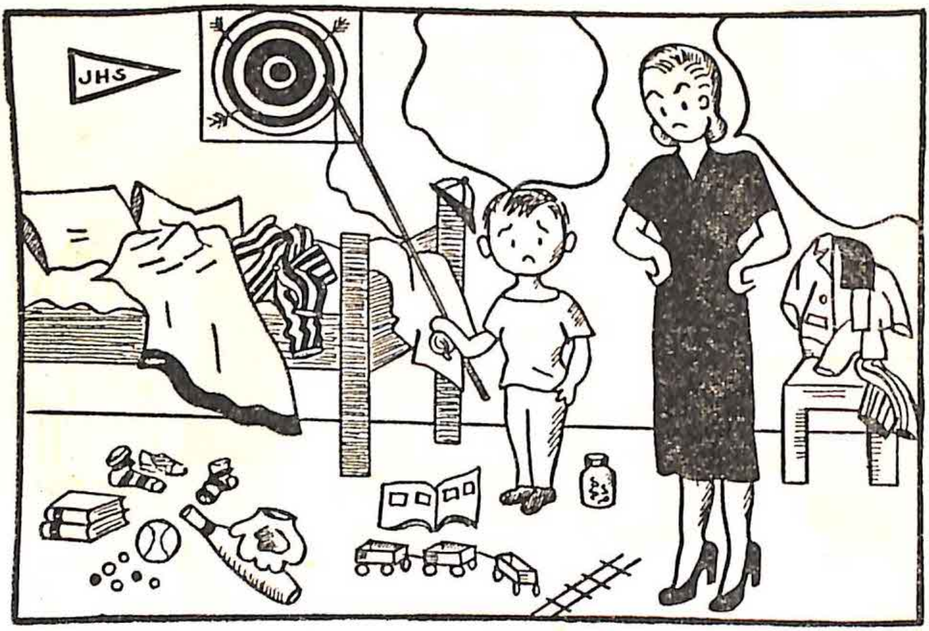


FIGURE 1—Card 6 of the Male Series of the Punishment Situation Index dealing with the lack of neatness in personal habits.

Card III. Socially unacceptable or socially embarrassing behavior. Male—Boy enters the room undressed while the mother is entertaining guests. Female—Girl enters the room undressed while the mother is entertaining guests.

Card IV. Interference by destruction of parents' personal possessions. Male—Boy has broken his mother's vase. Female—Girl dressing up in her mother's clothes and using her make-up.

Card V. Unfavorable intellectual attainment. Male—Boy has brought home a bad report card to his mother. Female—Girl has brought home a bad report card to her mother.

Card VI. Lack of neatness in personal habits. Male—Boy and his mother are shown in his room which is in a state of disorder. Female—Girl and her mother are shown in her room which is surrounded by disorder.

Card VII. Disobedience to parent. Male—Girl telling mother that the boy has disobeyed her by going swimming when told not to. Female—Boy telling mother that the girl has disobeyed her by going swimming when told not to.

Card VIII. Lying. Male—Girl telling mother that the boy is not telling the truth. Female—Boy telling mother that the girl is not telling the truth.

Card IX. Destruction of other's property. Male—Boy has broken a neighbor's window playing ball. Female—Girl picking a neighbor's flowers.

Card X. Stealing. Male—Girl telling mother that the boy took apples at the store. Female—Boy telling mother that the girl took candy at the store.

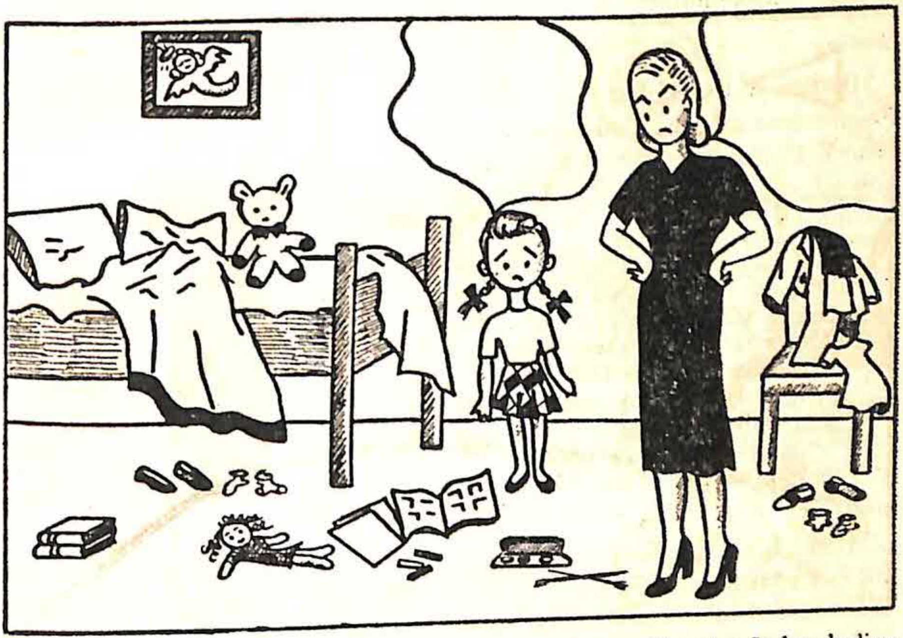


FIGURE 2—Card 6 of the Female Series of the Punishment Situation Index dealing with the lack of neatness in personal habits.

Pictorially, the cartoons are designed to be as unambiguous as possible to ensure that the subject will interpret them as having punishment implication. Spaces are provided above the figures as in comic strip cartoons for the subject to fill in what he thinks each character is saying. Because of the nature of the pictures, the PSI can be used with both mothers and children. When administered to both mothers and children, the PSI yields four concepts operating in the punishment situation: from the child, his self-concept (CC) and his concept of his mother (CM); from the mother, her self-concept (MM) and her concept of the child (MC). It was initially hypothesized that each subject, child or parent, would project his own attitudes toward punishment in his responses, and, where appropriate, would identify himself with the child or parent figure.

On the basis of exploratory administrations of the PSI, the following hypotheses are proposed:

I. Mothers are more extrapunitive in the punishment situation than their children, both according to themselves and their children.

II. Children are more intropunitive in the punishment situation than their mothers, both according to themselves and according to their mothers.

III. There is less impunitiveness in the punishment situation in both mother's and children's responses than either extrapunitiveness or intropunitiveness.

Scoring

Because of the similarity in design of the PSI to the *Rosenzweig Picture-Frustration* cards, the obtained responses were scored using the three scoring factors, Extrapunitiveness, Intropunitiveness and Impunitiveness for *direction of aggression,* developed by Rosenzweig (9, 10). Examples of the scoring of PSI responses are given below:

> *Extrapunitiveness* (E): Aggression is directed onto some person or thing in the environment. Typical of responses here are: (Child:) My mother didn't know what she was talking about. I felt mad at her. (Mother:) Will you *ever* learn to behave? You're a naughty little girl.
>
> *Intropunitiveness* (I): Aggression is directed by the subject onto himself. Typical of responses here are: (Child:) I shouldn't have done it. I need to be punished.
>
> *Impunitiveness* (M): Aggression is evaded in an attempt to gloss over the punishment. Typical of responses here are: (Mother:) That's just like a little girl. (Child:) I forgot it, I didn't even think of it anymore.

Combinations of any *two* of these three directions may occur within any one response. When two distinct thoughts were expressed, each response was counted as one-half point in the scoring scheme.

SUBJECTS AND PROCEDURE

Subjects for the present study were 24 fourth, fifth, and sixth grade children and their mothers. These children, made available for testing by the Highland Elementary School,[2] in Baton Rouge, Louisiana, ranged in age from 9 to 12 with a mean age for the group of 10.26. This age group was selected rather than a younger group because of their greater facility in speech and reading. The sample consisted of 14 girls and 10 boys, most of whom came from homes of professional fathers.

The Inquiry: Child

An inquiry was individually conducted in conjunction with each of the 10 pictures presented to the subject. The format of the inquiry for the child was as follows:

---

[2] The writers wish to express appreciation to Miss Geraldine Wall, Principal, for her helpful cooperation in making subjects available for this study.

Interviewer: "I am going to show you some pictures of a little boy (girl) and his (her) mother. In each picture, each boy (girl) is doing something different. I would like you to answer some questions about what they are doing."

The subject was then presented with a card, and the inquiry proceeded as follows:

Q-1. Tell me what is happening in this picture.
Q-2. Pretend you are the little boy (or girl) in this picture. Tell me what your mother would say if this happened. (CM)
Q-3. What would you say to your mother? (CC)
Q-4. What would your mother do if this happened? (CM)
Q-5. What would you do if this happened? (CC)
Q-6. How would you feel when your mother did that? (CC)
Q-7. How do you think your mother would feel when she did that? (CM)

The Inquiry: Mother

The interviewer explained to the mother at the outset of the session: "In the attempt to arrive at a better understanding of the relationships between mothers and their children, we are currently investigating the ways in which parents teach their children things such as honesty, obedience, and so forth."[3]

Since the mother might refer to a child other than the subject selected for the study, the interviewer was instructed always to make sure he knew which child the subject was referring to, especially when she referred to him by name.

In the case of mothers who had two children of the same sex and approximately the same age, the interviewer instructed the parent that she was to consider her answers in terms of the one child only. To ensure that the subject did this, the interviewer asked the mother to mention the name of her child in each question where the word *child* appeared. The following instructions were then given: "We have some pictures of children and their mother in typical situations that most parents have to deal with at some time. I would like you to answer some questions about what they are doing in the pictures."

Q-1. What do you think has happened in this picture?
Q-2. What do you think you would say if this were your child? (MM)
Q-3. What do you think your child would say? (MC)
Q-4. What would you do if it happened? (MM)
Q-5. What would your child do if it happened? (MC)
Q-6. How do you think he would feel when you did that? (MC)
Q-7. How would you feel at the time you did that? (MM)

---

[3] It was obviously anticipated that the mothers would exhibit greater curiosity than the children as to the purpose of the study. The preliminary explanation to the mothers was used to establish rapport. Since the actual purpose was not revealed to either the mothers or the children, it is felt that the instructions did not differentially affect their responses.

RESULTS AND DISCUSSION

Each of the inquiries was scored according to the method developed by Rosenzweig (2, 9, 10). Responses to Questions 2, 4, and 7 of the children's inquiry were concerned with the child's concept of his mother (CM), and Questions 3, 5, and 6 dealt with the child's concept of himself (CC). Questions 3, 5, and 6 of the mother's inquiry dealt with the mother's concept of the child (MC), and Questions 2, 4, and 7 of this inquiry were concerned with the mother's self-concept (MM). The mean number of times each method of direction of aggression was employed by either the child or the mother will be considered here in relation to testing each hypothesis outlined.

TABLE 1

COMPARISON OF DIFFERENCES IN EXTRAPUNITIVENESS (E) ATTRIBUTED TO MOTHERS AND CHILDREN

Variables					
1		2	$Mean_1$	$Mean_2$	t
CM_E	$>$	MC_E	22.00	8.77	12.88*
MM_E	$>$	CC_E	17.42	7.90	8.21*
MM_E	$>$	MC_E	17.42	8.77	8.32*
CM_E	$>$	CC_E	22.00	7.90	12.21*
CM_E	$>$	MM_E	22.00	17.42	3.98*
MC_E	$>$	CC_E	8.77	7.90	.83

* Significant beyond the .01 level.

Hypothesis I

According to this hypothesis, in the punishment situation mothers are more extrapunitive than their children, both according to themselves and their children. If this relationship is found to hold, we would predict that $CM_E > MC_E$, $MM_E > CC_E$, $MM_E > MC_E$, and $CM_E > CC_E$. To test this relationship, t tests for matched groups on extrapunitiveness scores were computed. As shown in Table 1, the differences were not only in the expected direction, but in every case, were significant beyond the 1 per cent level. These findings indicated that the mothers tend to project blame onto another person and/or the environment from both their own and their children's points of view.

From the means contained in Table 1, it appears that the children's concept of the mother's extrapunitiveness is considerably greater than the mother's concept of her own outward aggression, and likewise, the mother's conception of the children's aggression greater than the child's self-concept. These subhypotheses were tested using a t test of significance. The theory that children conceive of mothers as more extrapunitive than mothers conceive themselves to be was clearly sustained, and is in line with the psychoanalytic contention that the parent figure is regarded as threatening by the child. The idea that the mother has a greater conception of the child's extrapunitiveness, however, was rejected.

Hypothesis II

According to the second hypothesis, children are more intropunitive in the punishment situation than their mothers, both according to themselves and their mothers. For this hypothesis to be sustained, it is necessary for the following relationship to hold: $MC_I > CM_I$, $CC_I > MM_I$, $MC_I > MM_I$, and $CC_I > CM_I$. The results of the t tests applied here are shown in Table 2. In all cases, significance exceeded the 1 per cent level, leading to the conclusion that children are more intropunitive than their mothers both according to themselves and their mothers.

TABLE 2

COMPARISON OF DIFFERENCES IN INTROPUNITIVENESS (I) ATTRIBUTED TO MOTHERS AND CHILDREN

Variables					
1		2	$Mean_1$	$Mean_2$	t
CM_I	<	MC_I	1.52	10.17	11.98*
MM_I	<	CC_I	5.71	10.08	3.83*
MM_I	<	MC_I	5.71	10.17	5.41*
CM_I	<	CC_I	1.52	10.08	8.02*
CM_I	<	MM_I	1.52	5.71	7.16*
MC_I	>	CC_I	10.17	10.08	.07

* Significant beyond the .01 level.

Inspection of the mean frequencies also indicated the relationship: $MM_I > CM_I$, and $CC_I > MC_I$. That mothers conceive of themselves as more intropunitive than children conceive them to be was clearly indicated by the t test; that children conceive of themselves as more intropunitive than mothers conceive them to be, however, failed to meet significance.

As in the case for the extrapunitive expression of aggression, we can conclude that the mother's concept of the child is in close agreement with the child's concept of his intropunitiveness. This finding confirms the view suggested above: that children regard the parent-figure as threatening, and as a result, tend to overestimate the mother as a punishing agent (extrapunitiveness), and underestimate her as intropunitive. That is, children believe that the mother does not often take the blame herself for incidents which occur, but rather tends to blame them.

Hypothesis III

In line with our third hypothesis, we were interested in determining whether impunitiveness occurs less frequently in punishment situations than either extrapunitiveness or intropunitiveness. If this is to be sustained, then all measures of E and I for both the mothers and children must be greater than all measures of M. Since the frequency of M responses was too small to give a normal distribution, the sign test was here employed (see Table 3). This third hypothesis was shown to hold at the 1 per cent level in all but two cases: $CC_I - CC_M$ was found significant at the 5 per cent level; $CC_E - CC_M$ was found nonsignificant.

TABLE 3

RESULTS OF SIGN TEST COMPARING EXTRAPUNITIVENESS, INTROPUNITIVENESS, AND IMPUNITIVENESS ATTRIBUTED TO MOTHERS AND CHILDREN

Variables	*N*[*]	*r*[†]	*Level of Significance*
$CM_E > CM_M$	24	0	.01
$MM_M > MM_M$	24	0	.01
$MC_E > MC_M$	22	3	.01
$CC_E > CC_M$	24	8	.25
$CM_I > CM_M$	20	3	.01
$MM_I > MM_M$	22	3	.01
$MC_I > MC_M$	24	1	.01
$CC_I > CC_M$	21	5	.05

[*] *N* is the number of persons in the sample, ignoring all pairs with a difference of zero.
[†] *r* is the smaller of two values.

It also appeared that more impunitiveness was attributed to children than to the mother group; that is, the following relationship should be sustained: $MC_M > CM_M$, $CC_M > MM_M$, $MC_M > MM_M$, and $CC_M > CM_M$. The application of the sign test again sustained these relationships (Table 4) at beyond the .01 level of confidence, supporting the hypothesis that children are more impunitive than mothers, both according to themselves and according to their mothers. From these findings, it appears that children in the punishment situation tend to blame themselves more than mothers. Children also appear to evade or gloss over frustration more than their mothers. As a result, mothers direct more blame and hostility onto the children than the children direct towards their mothers. And children appear to overestimate the amount of blame directed towards them in comparison with what their mothers believe they do. In line with their overestimation of mother's extrapunitiveness, children also underestimate the amount of blame mother's take themselves as compared to what mothers believe they do.

TABLE 4

COMPARISON OF DIFFERENCES IN IMPUNITIVENESS (M) ATTRIBUTED TO MOTHERS AND CHILDREN

Variables	*N*	*r*	*Level of Significance*
$MC_M > CM_M$	23	0	.01
$CC_M > MM_M$	23	4	.01
$MC_M > MM_M$	21	0	.01
$CC_M > CM_M$	22	0	.01

SUMMARY AND CONCLUSIONS

The development of a new projective test—the *Picture Situations Index*—was undertaken in an attempt to investigate and assess the direction of aggression in the mother-child punishment situation. Two sets of 10 cards each, one for boys and one for girls, were designed to depict common situations often leading to punishment. This test was administered to 24 school children between the ages of 9 and 12, and their mothers. The inquiry for each card, a rather structured cartoon-like sketch, was designed to yield four concepts usually operative in the punishment situation: the child's concept of the mother (CM); the child's self-concept (CC); the mother's concept of the child (MC); and the mother's self-concept (MM).

Three measures of the direction of aggression, Extrapunitiveness, Intropunitiveness, and Impunitiveness, were utilized in the data analysis. On the basis of the emerging concepts of the mother and child obtained through this scoring method, the following conclusions appear warranted: (*a*) Mothers direct more aggression onto their children than children direct onto mothers. (*b*) Children conceive of mothers as directing more aggression onto them than mothers conceive of themselves as doing. (*c*) Children direct more of their aggression onto themselves than mothers direct onto themselves. (*d*) Mothers conceive of themselves as directing more aggression onto themselves than children conceive of them as doing. (*e*) Children are more likely than mothers to gloss over and evade punishment.

The major conclusion warranted from the present results is that the merit of the PSI for studying the mother-child situation through the use of "concepts" lies not in the "validity" of the responses as such, but in the finding that mother and child view it from a different frame of reference, as evidenced in the discrepancies of their responses to the same objective stimulus card. The discrepancies between self-perception and others' perception of the self could result in the difference between mutual understanding and perceived rejection leading to possible punishment, unwarranted or otherwise. If we can assume that both child and mother reveal significant aspects of personality in their responses, the PSI might provide the much needed "bridge" between punishment and personality which hithertofore has been dismissed by reasoning of the type: this child is delinquent, he was severely punished, ergo, he is delinquent because he was severely punished. Such a technique as has been described here might direct light on some of the more obscure aspects of the punishment situation.

REFERENCES

1. AMMONS, CAROL H., & AMMONS, R. B. Aggression in doll-play: interviews of two-to-six-year-old white males. *J. genet. Psychol.*, 1953, 82, 205-213.
2. CLARKE, HELEN JANE, ROSENZWEIG, S., & FLEMING, EDITH. The reliability of the scoring of the Rosenzweig P-F study. *J. clin. Psychol.*, 1947, 3, 364-470.
3. DOLLARD, J., MILLER, N. E., DOOB, L. W., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.
4. HAMMER, E. F. Frustration-aggression hypotheses extended to socio-racial areas: comparison of Negro and white children's H-T-P's. *Psychiat. Quart.*, 1953, 27, 597-607.
5. HOFFMAN, M. L. Some psychodynamic factors in compulsive conformity. *J. abnorm. soc. Psychol.*, 1953, 48, 383-393.
6. JERSILD, A. T., & MARKEY, F. V. Conflicts between preschool children. *Child Develpm. Monogr.*, 1935, No. 21.
7. MOWRER, O. H., & KLUCKHOHN, C. Dynamic theory of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald Press, 1944, Vol. I., pp. 69-135.
8. REINHARDT, J. M., & FOWLER, V. H. Social and ethical judgments of two groups of boys—delinquents and non-delinquents. *Amer. J. Law Criminol.*, 1930, 21, 364-378.
9. ROSENZWEIG, S., CLARKE, HELEN J., GARFIELD, MARJORIE S., & LEHNDORFF, ANNEMARIE. Scoring samples for the Rosenzweig P-F study. *J. Psychol.*, 1946, 21, 45-72.
10. ROSENZWEIG, S., FLEMING, EDITH, & CLARKE, HELEN J. Revised scoring manual for the Rosenzweig P-F study. *J. Psychol.*, 1947, 24, 165-208.

11. Sears, R. R. Non-aggressive reactions to frustration. *Psychol. Rev.*, 1941, 48, 343-346.

12. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, Pauline S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-236.

13. Symonds, P. M. *The psychology of parent-child relationships.* New York: D. Appleton-Century, 1939.

CALCIFICATION OF THE MANDIBULAR THIRD MOLAR AND ITS RELATION TO SKELETAL AND CHRONOLOGICAL AGE IN CHILDREN[1]

ARTHUR DEMISCH[2] *and* PETER WARTMANN[3]
Orthodontic Department, Forsyth Dental Infirmary for Children

Whereas theoretically it may be expected that an intimate relationship exists in the physical maturation of various tissue systems and that this relationship is reflected by the different maturity indicators, the correlation between dental development and growth of the body as a whole is claimed to be low (1, 8).

Yet Talmers (9) has demonstrated that children with late emergence of permanent teeth were usually delayed in both height and weight gain when appraised by referring to the Fels norms, and that children with early eruption of teeth were advanced in their physical growth. The asso-

---

[1] The authors thank Dr. C. F. A. Moorrees, Chief of the Orthodontic Department, for his guidance and counsel; Dr. V. O. Hurme, Director of Clinical Research, for critical review of the manuscript; and Miss Marguerite Prioli, R.D.H., also of the Forsyth Dental Infirmary for Children, for taking the radiographs.

The authors gratefully acknowledge the valuable advice received from two staff members of the School of Public Health, Harvard University; namely, Dr. S. Idell Pyle, Consultant in Developmental Anatomy, and Dr. Robert B. Reed, Associate Professor of Biostatistics and Human Ecology.

[2] Intern at the Forsyth Dental Infirmary for Children (1953-1954). At present First Assistant in Orthodontics at the University of Bern Dental School, Switzerland.

[3] Intern at the Forsyth Dental Infirmary for Children (1953-1954). Research Fellow in the Department of Periodontology at the University of Zürich Dental School, Switzerland (1954-1955).

CHILD DEVELOPMENT, Vol. 27, No. 4 (December, 1956)

ciation between these three variables was demonstrated most clearly for children with late eruption of teeth.

Gleiser and Hunt (2, 4) studied the chronology of calcification of the permanent mandibular first molar in considerable detail and reported that arrests in the ossification of hand and wrist bones often coincide with arrests in tooth formation.

The present study deals with the calcification of the mandibular third molar since it is desirable, especially for the orthodontist, to have this information for evaluating dental development between 9 and 15 years of age, or after the permanent mandibular first molar can be used for this purpose. Such data can complement the estimate of dental maturation based on the clinical emergence of permanent teeth, and Gleiser and Hunt (2) suggest "that the calcification of a tooth may be a more meaningful indication of somatic maturation than is its clinical emergence." The relationships between third molar development and chronological and skeletal age were studied also.

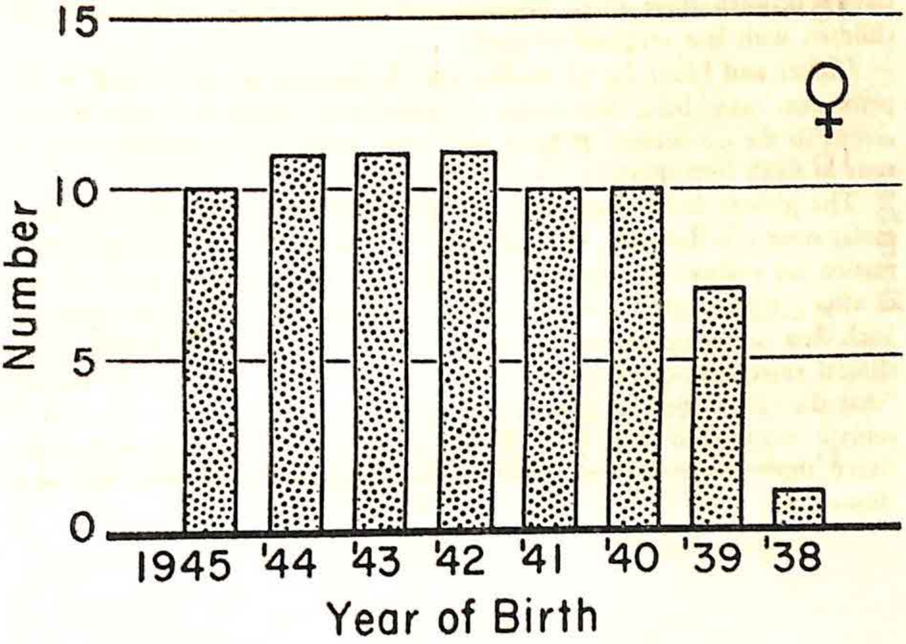
THE PRESENT STUDY

A total of 151 American white children of Greater Boston, who were patients at the Forsyth Dental Infirmary for Children and without obvious signs of nondental disease or of developmental disturbances, were selected for examination according to the year of birth (Fig. 1). When, however, the 81 boys and 70 girls are grouped according to age (Fig. 2), the resulting frequency distributions differ considerably from those based on the year of birth, due to the time lapse between the examination of the first and last subject. In each instance, almost invariably on the same visit, radiographs of the right hand and wrist as well as of the jaw were taken.

The hand radiographs were analyzed according to the technique of S. Idell Pyle by computing the mean skeletal age after comparing each bone of the hand and wrist to the standards of Greulich and Pyle (3). Since in this study only one evaluation of skeletal age was needed, rather than a longitudinal picture of skeletal development, no consideration was given to the differences in maturation of the various hand bones of the individual child, as required by the "Red Graph" method (6).

Differences in evaluating skeletal age were determined by independent examinations of the hand radiographs of 30 children (15 boys and 15 girls). The results, summarized in Table 1, show a high degree of similarity between the readings made by the two observers.

For the evaluation of the calcification of the mandibular third molars, use was made of the standards developed by Gleiser and Hunt for the first molar. Yet, independently of the authors just mentioned, an attempt was made in the initial stage of this study to find suitable developmental stages



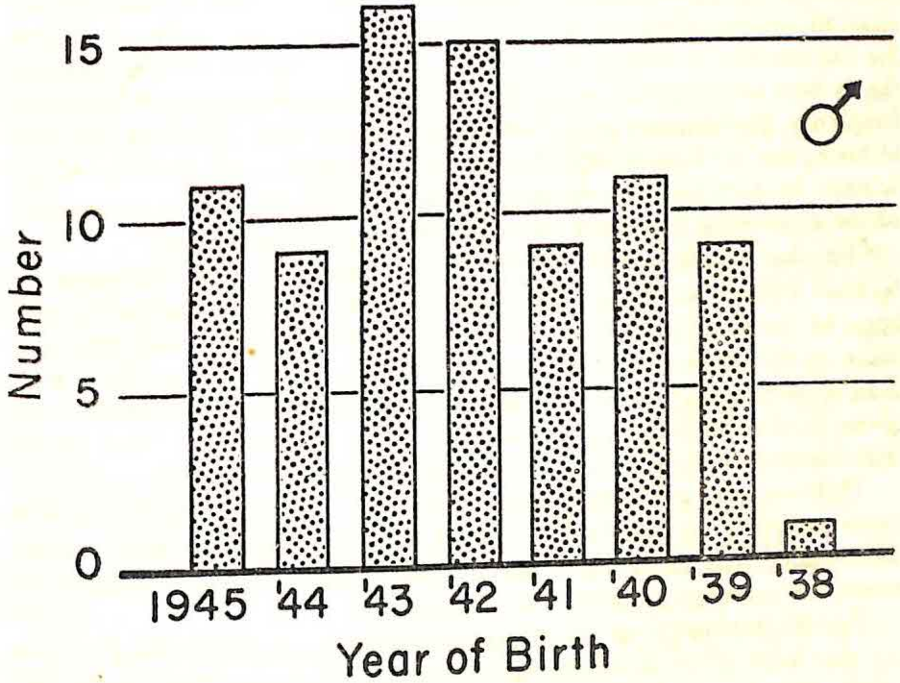


FIGURE 1—The frequency distributions according to the year of birth of the 70 girls and 81 boys studied.

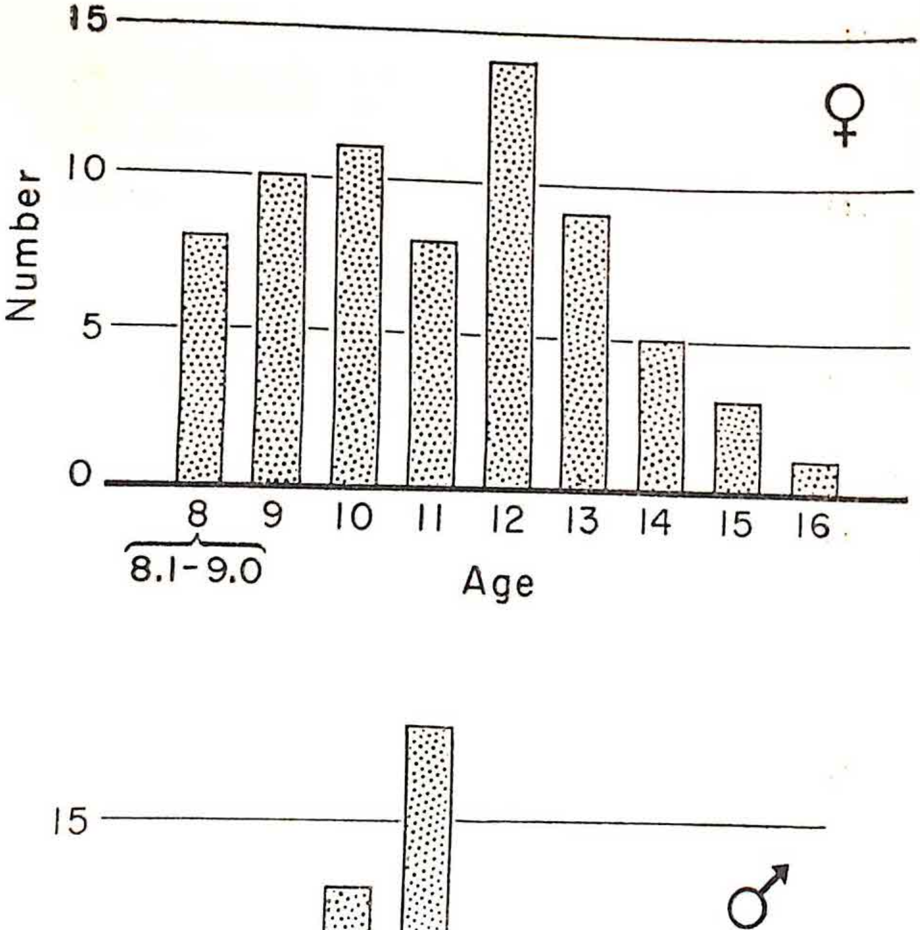


FIGURE 2—The frequency distributions according to the chronological age of the 70 girls and 81 boys studied.

TABLE 1

DIFFERENCES BETWEEN INDEPENDENT EVALUATIONS OF A SERIES OF 30 HAND RADIOGRAPHS BY TWO OBSERVERS

Rating D > Rating W			*Rating D = Rating W*	*Rating D < Rating W*		
9 mo.	6 mo.	3 mo.		3 mo.	6 mo.	9 mo.
0	2	2	22	1	3	0

for gauging the calcification of the third molar. This resulted in almost the same scale as that of Gleiser and Hunt, except that stage 6—2/3 of crown completed—could not be differentiated accurately from its neighboring stages and it was, therefore, omitted. Stage 6 in the present study, signifying completion of the crown, thus corresponds to stage 7 of Gleiser and Hunt and stage 7, the beginning of root formation, to their stage 8. The upper age limit of the sample was not high enough to study the subsequent stages of root formation until the terminal convergence of the root canal. The definitions of each calcification stage as used in this study are given in Table 2.

Unfortunately, it was not possible to obtain satisfactory roentgenologic records for classifying the development of the maxillary third molars, which explains why this investigation was limited to studying the mandibular molars.

Differences in assessing mandibular third molar development were determined by independent evaluations of 50 lateral jaw radiographs. Since

TABLE 2

DEFINITIONS OF EACH STAGE OF THIRD MOLAR CALCIFICATION USED IN THIS STUDY

0	No change in bone density, and no crypt visible.
1	Crypt clearly visible, but no calcification.
2	Calcification of the tips of one to four cusps.
3	Coalescence of two or more centers.
4	Outline of the cusps completed, calcification progressing towards the level of the grooves, but the center of the occlusal crown surface may not yet be calcified.
5	Half of the crown completed, calcification up to the largest mesiodistal diameter of the tooth crown, but the entire enamel formation not yet completed.
6	Crown completed, enamel formation completed to enamel-cementum junction.
7	Beginning of root formation.

stage 5 offered the greatest difficulties for proper classification, a larger number of teeth in this stage than in the other stages of development was selected by the first examiner to test whether his co-observer agreed with the rating given. The disagreement in evaluations of third molar development was never more than one stage, and the range of identical ratings varied between 60 and 100 per cent, as shown in Table 3.

TABLE 3

DIFFERENCES BETWEEN INDEPENDENT EVALUATIONS OF 50 RADIOGRAPHS OF THE MANDIBULAR THIRD MOLAR BY TWO OBSERVERS

	STAGE OF CALCIFICATION								
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*Total*
Number of radiographs	5	6	4	5	6	14	5	5	50
Rating D = Rating W	5	4	4	5	4	10	3	4	39
Rating D ≠ Rating W	0	2	0	0	2	4	2	1	11
Per cent of identical ratings . . .	100	67	100	100	67	71	60	80	78

The reliability of evaluations of both the wrist and lateral jaw radiographs is considered adequate in view of the following reasons: (1) In general, no tendency for consistent over- or underrating by one observer compared to the other was observed. (2) In testing for the differences when two observers independently studied a series of hand radiographs, the same mean bone age was found in 73 per cent of trials, and, when testing the evaluations of molar calcification, identical ratings were obtained in 78 per cent of the double determinations. The number of identical ratings is thus similar in both instances.

Differences in evaluating third molar development do not exceed one stage or one-sixth of the entire scale of crown formation. The differences in the estimate of bone age are no greater than 6 months which corresponds, according to Gleiser and Hunt, to about one-seventh of the time needed for the formation of the crown of the first molar. If it is assumed, for practical purposes, that the speed of formation of the mandibular third molar crown is approximately the same as that of the first molar, the differences in independent evaluations of molar development and bone age do not exceed one-sixth of the time needed for completion of the third molar crown, or 7 months in the boys and 9.7 months in the girls. However, for bone age only the maximal differences amount to 6 months, while for molar calcification the differences can never be less than one stage. Therefore, the errors made by the two observers in studying third molar development are often greater than those made by them when assessing skeletal age.

FINDINGS

The findings can be grouped under four headings, namely: (1) the relationship between skeletal age and chronological age, (2) degrees of bilateralism in the development of left and right third molars in the mandible, (3) the median chronological and skeletal ages for each calcification stage of these third molars, (4) the mutual relationships between third molar calcification and skeletal and chronological age.

1. The mean difference between skeletal age and chronological age is zero months in the 81 boys and 6 months in 70 girls studied (Table 4). Thus in the average girl the skeletal age is one-half year advanced compared to the chronological age. The range of approximately 48 months for both sexes corresponds closely to the findings of Greulich and Pyle. In a group of "normal" early and late maturing children one may, therefore, expect differences as great as 24 months between the two measures of development.

TABLE 4

THE MEAN DIFFERENCE AND STANDARD DEVIATION OF SKELETAL AGE AND CHRONOLOGICAL AGE

	N	*Mean*	*S.D.*
Boys	81	0 months	± 12.3 mo.
Girls	70	6 months	± 10.1 mo.

2. Chronological symmetry in the development of the mandibular left and right third molars occurs in 71 per cent of the boys and girls. In the others differences of one, but not more than two, developmental stages are observed in the calcification of antimeres. It is not possible to conclude from the data at hand whether a definite trend exists for more advanced development on one side of the mouth than on the other.

Saito (7) reported earlier calcification and eruption of third molars on the right side of the mouth when compared with those on the left, and he also found that Japanese girls had a more rapid developmental rate than Japanese boys until the age of 13 years. This last conclusion is not corroborated by the findings of the present study.

Unilateral agenesis must be considered also as an asymmetry of development. However, great care was necessary in diagnosing agenesis in this group of young Bostonians inasmuch as the beginning of calcification, as manifested in radiographs (stage 1), was observed as late as 12 years and 5 months in some children.

TABLE 5

AGENESIS OF MANDIBULAR THIRD MOLARS IN CHILDREN "AGED 13.1 YEARS AND OLDER"

	N	*Unilateral*	*Bilateral*	%
Boys	21	1	1	9.5
Girls	17	2	4	35.3

In the small sample of boys 13 years 1 month and older (Table 5), the combined frequency of unilateral and bilateral mandibular third molar agenesis was 9.5 per cent and in the still smaller group of girls it was as high as 35.3 per cent. After an analysis of extensive radiographic material Saito (7) reported "congenitally missing" mandibular third molars in 20 per cent of his Japanese sample, while Nanda (5) found agenesis of one [illegible]

TABLE 6

MEDIAN CHRONOLOGICAL AND SKELETAL AGES AS WELL AS AGE RANGES OF THE DIFFERENT STAGES OF MANDIBULAR THIRD MOLAR CALCIFICATION

		Median		Range	
M_3	*N*	*CA*	*SA*	*CA*	*SA*
		BOYS	(*N* = 69)		
1	8	9.2	9.3	8.5 –12.0	8.0 –11.0
2	6	10.3	10.6	9.6 –11.5	9.0 –11.6
3	16	11.1	10.6	8.9 –13.0	8.0 –12.0
4	12	11.7	11.9	10.4 –13.7	10.0 –13.0
5	9	12.11	13.3	10.5 –13.9	11.6 –14.0
6	7	14.0	14.0	10.9 –15.2	12.0 –16.2
7	11	14.7	15.0	12.11–16.1	14.0 –17.0
		GIRLS	(*N* = 51)		
1	5	9.3	9.0	8.6 –12.5	7.6 –11.6
2	3	10.3	10.0	9.9 –11.4	8.0 –12.0
3	9	10.5	11.0	9.0 –14.9	9.0 –13.0
4	13	11.5	11.6	8.11–12.11	9.0 –13.0
5	9	13.6	14.0	11.11–14.6	12.0 –14.2
6	3	12.9	13.0	11.10–12.10	10.0 –13.6
7	9	14.11	15.0	12.4 –15.4	13.6 –16.6

NOTE.—Units in years, fractions in months, CA = chronological age, SA = skeletal age, M_3 (1) = stage of calcification.

or more third molars—maxillary or mandibular—in 9 per cent of 216 North American white women.

Unilateral agenesis of the mandibular third molars did not change the statistical findings concerning symmetry of development because it occurred in only three instances.

3. The median age and the age ranges of each of the seven calcification stages studied are given in Table 6. For these statistics an average calcification rating was used in the case of asymmetric development of antimeres, except that when in an individual one mandibular third molar was rated as being in stage 0, the higher rating of its antimere was used as the overall rating. In the majority of subjects (71 per cent of the total sample), however, the calcification of the left and right third molars was symmetrical, as pointed out already.

In both sexes the median chronological and skeletal ages of third molar stages agree closely, the maximum difference being 7 months for stage 3 in both the males and females.

The numbers of mandibular third molars in each calcification stage in each sex and in the two sexes combined are presented in Table 7, together with the percentage frequencies of teeth in each stage.

TABLE 7

THE NUMBER OF BOYS AND GIRLS WITH MANDIBULAR THIRD MOLARS IN DIFFERENT STAGES OF CALCIFICATION AND THE PERCENTAGE FREQUENCIES OF THIRD MOLARS IN EACH STAGE

(Total number of children = 151, total number of teeth = 302; agenesis not considered)

	STAGE OF CALCIFICATION							
	0	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Boys	33	14	6	28	25	17	14	25
Girls	46	5	4	17	24	19	6	19
Boys + Girls	79	19	10	45	49	36	20	44
Per cent of all teeth	26	6	3	15	16	12	7	15

4. An important aspect of this investigation was to determine whether individuals with early or late skeletal age have a correspondingly advanced or delayed developmental pattern of their mandibular third molars. This relationship could be studied in 69 boys and 51 girls only since in some subjects the first sign of calcification of the mandibular third molars was not seen on the radiographs.

Scatter diagrams were prepared for preliminary analysis of the correlation between the different variables, an average calcification rating being used in case of asymmetric development of mandibular left and right third molars. Frequency distributions of the observations are given on the *X* and *Y* axes of each scattergram.

After studying the scattergrams (Figs. 3 and 4) it was concluded that the regression lines follow a straight trend. However, when the median chronological and skeletal ages for each developmental stage are plotted on the scattergrams, the points are not found to be on a straight line; but their arrangement is such that it seems justified to assume a straight line trend. Therefore, Pearson's product-moment correlation formula was considered satisfactory for further statistical treatment of the material.

For computations of the correlation coefficients fractions of class intervals of molar calcification, resulting from averaging the findings for left and right third molars, were reduced to full stages (e.g., a rating of 3.5 was included in the stage 3 group).

According to the findings (Table 8), a high degree of association exists between third molar calcification and skeletal or chronological

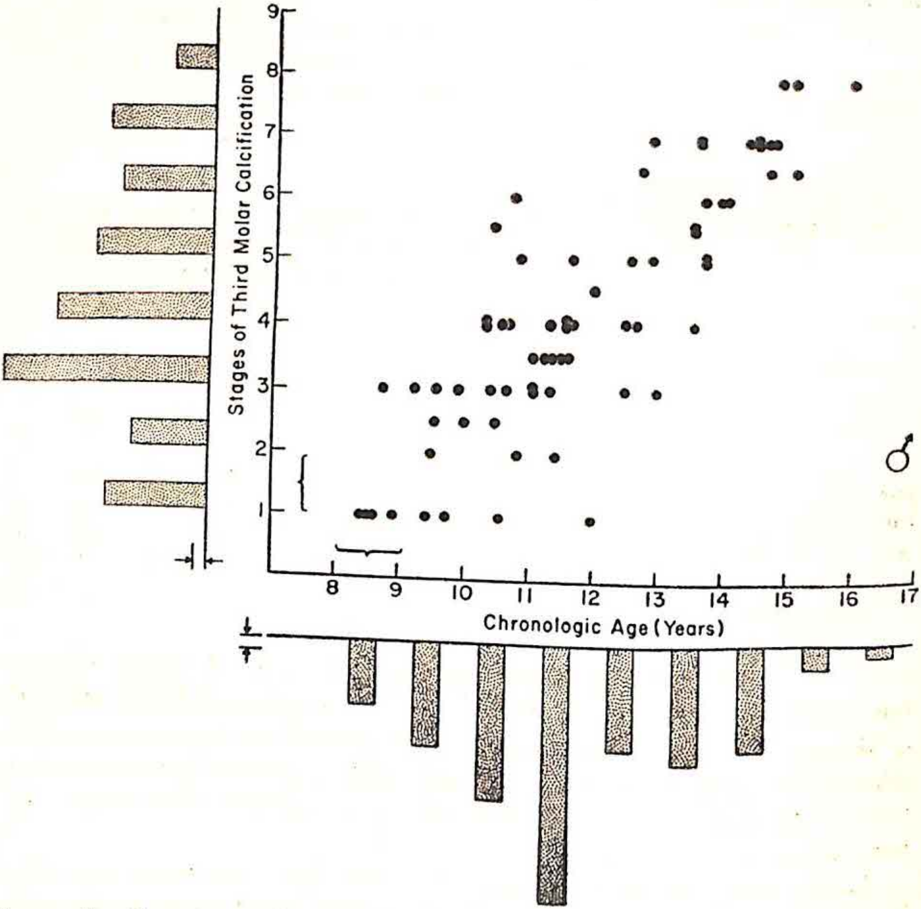


FIGURE 3a—Scattergram to demonstrate the degree of association between stages of third molar calcification and chronological age for 69 boys.

TABLE 8

COEFFICIENTS OF CORRELATION (r) AND THE STANDARD ERRORS ($S.E._r$) OF MANDIBULAR THIRD MOLAR CALCIFICATION (M_3), CHRONOLOGICAL AGE (CA) AND SKELETAL AGE (SA), AND THE MULTIPLE CORRELATION COEFFICIENTS ($R \pm S.E._R$) BETWEEN THESE THREE VARIABLES

	Boys $N = 69$		*Girls* $N = 51$		*Boys and Girls* $N = 120$	
	r	$S.E._r$	r	$S.E._r$	r	$S.E._r$
M_3 : CA	.83	± .037	.73	± .064	.79	± .034
M_3 : SA	.86	± .031	.75	± .060	.83	± .028
CA : SA	.89	± .024	.92	± .021	.90	± .016
	R	$S.E._R$	R	$S.E._R$	R	$S.E._R$
SA : CA M_3	.92	± .019	.93	± .019	.92	± .014

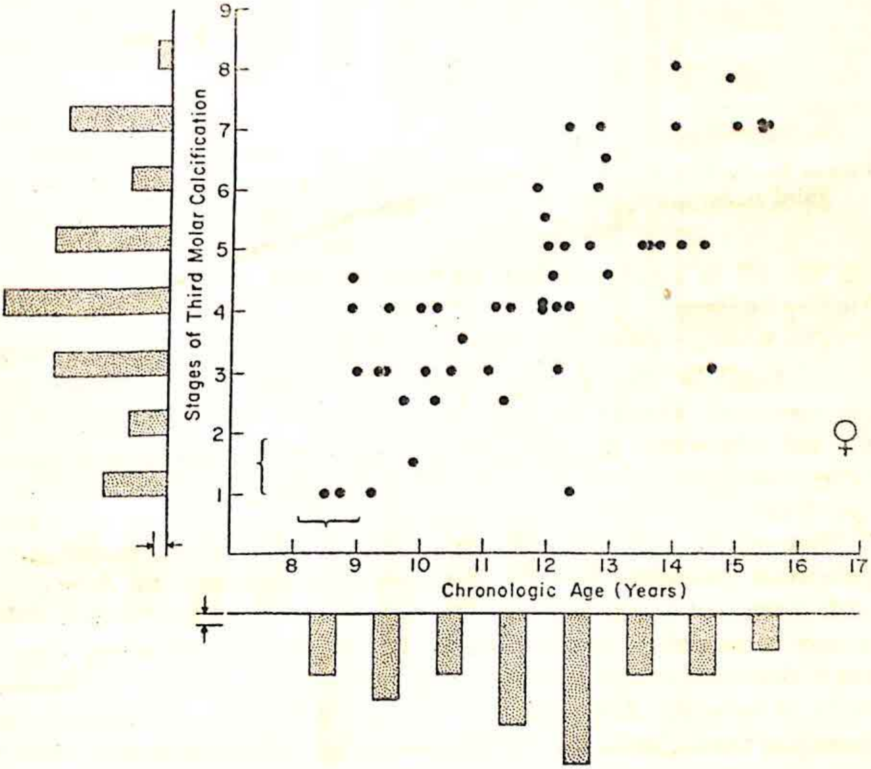


FIGURE 3b—Scattergram to demonstrate the degree of association between stages of third molar calcification and chronological age for 51 girls.

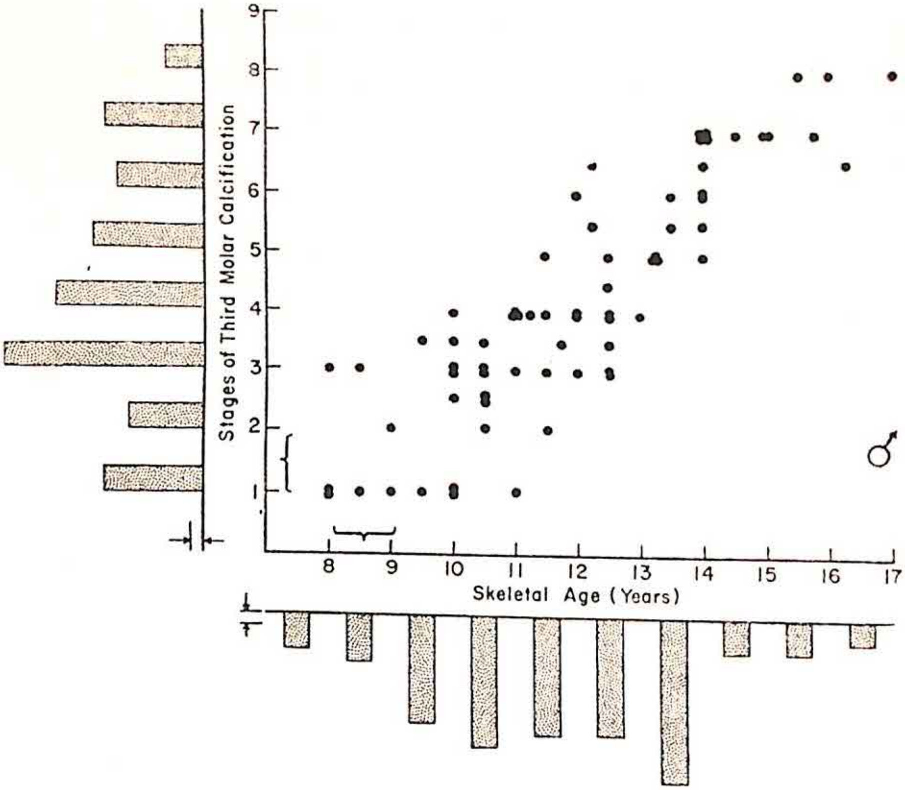


FIGURE 4a—Scattergram to demonstrate the degree of association between stages of third molar calcification and skeletal age for 69 boys.

age as well as between skeletal and chronological age (with values of r ranging between $+.73$ and $+.92$).

The multiple correlation coefficient is likewise high ($R_{SA.CAM_3}$ being $+.92$ or $.93$), but these findings do not differ greatly from those measuring the association between chronological and skeletal age. Therefore, the additional information on third molar calcification enables only a slightly better estimate of skeletal age than that obtained by the use of chronological age alone.

Nevertheless, in view of the comparatively small but significant partial correlation coefficients between third molar calcification and skeletal age with chronological age held constant ($r_{MSA.CA} = +.45$ for boys and girls), it may be possible to estimate skeletal age indirectly when taking into account the development of other permanent teeth in addition to the third molar if adequate calcification data for these other teeth are available, since averaging the variations in the development of individual teeth is likely to give one a more realistic evaluation of the dental age. If the variation in the calcification of different teeth is extensive, it may turn out that a reliable

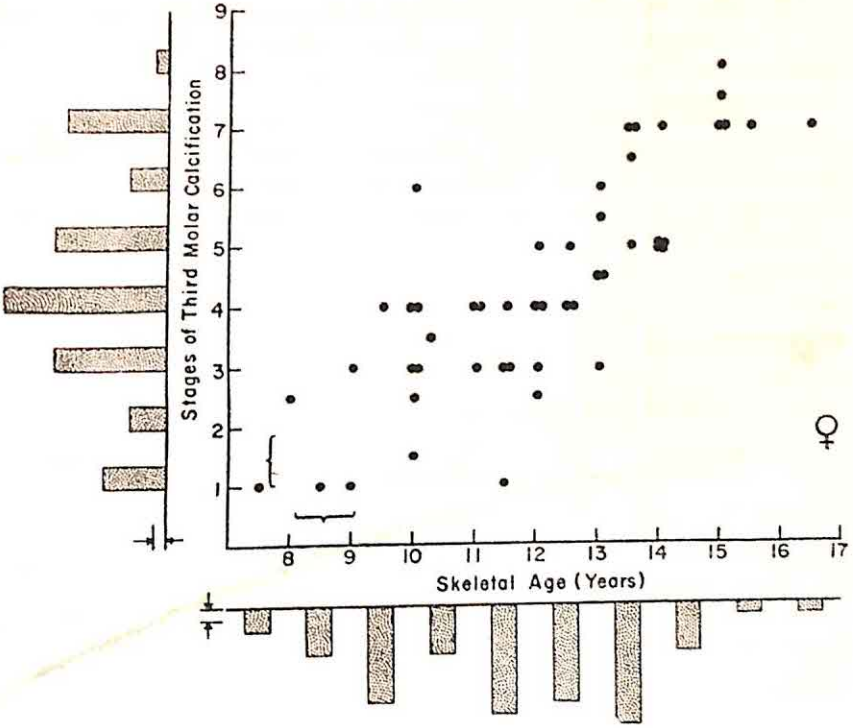


FIGURE 4b—Scattergram to demonstrate the degree of association between stages of third molar calcification and skeletal age for 51 girls.

estimate of skeletal age cannot be made and in such instances it would be of interest to find out from the actual assessment of bone age, by means of the Red Graph, whether the development of the hand and wrist bones also shows a wide distribution around the mean skeletal age. In addition, such efforts would show whether third molar calcification is more variable in comparison to that of other permanent mandibular teeth.

As a pilot study the data presented support the basic theoretical contention that a relationship exists in the physical maturation of various tissue systems, and the chronology of dental development may prove to be useful as an indicator of physiologic age. Exactly how valuable it is for this purpose remains to be established, preferably by a study of longitudinal data on a sufficiently large group of individuals.

SUMMARY

1. Radiographs of the right hand and each half of the jaws of 151 North American white individuals aged 8 to 16 years were studied for assessment of skeletal age and for the development of the mandibular third molars.

For estimating skeletal age, the standards of Greulich and Pyle were utilized and for molar calcification, the (slightly modified) standards of Gleiser and Hunt were used.

2. The results confirm the earlier studies concerning the wide range of skeletal ages compared with the chronologic ages in "normal" children. The differences when two observers evaluate the same wrist radiographs independently are much smaller than the "normal" variation of skeletal age in "healthy" children at a given chronological age. The differences in independent ratings of third molar development are somewhat higher but still adequate for the purpose of this study. Gleiser and Hunt's scale for rating the calcification of mandibular molar crowns appears satisfactory for clinical studies, except that their stage 6, which could not be differentiated clearly from its neighboring stages, had to be omitted.

3. The calcification of the mandibular third molar is symmetrical in 71 per cent of the children studied. In the others, no tendency is seen for consistently advanced development of either the left or right teeth.

4. There is a high positive correlation, with an approximately straight line trend, between the degree of calcification of the mandibular third molar and the skeletal or chronological ages. The findings support the theoretical contention that a relationship exists between the maturation of various tissue systems.

REFERENCES

1. CATTELL, P. *Dentition as a measure of maturity.* Cambridge: Harvard Univer. Press, 1928.
2. GLEISER, I., & HUNT, E. E., JR. The permanent mandibular first molar: its calcification, eruption and decay. *American J. phys. Anthrop.,* 1955, 13, 253-284.
3. GREULICH, W. W., & PYLE, S. I. *Radiographic atlas of skeletal development of the hand and wrist.* Stanford, California: Stanford Univer. Press, 1950.
4. HUNT, E. E., JR., & GLEISER, I. The estimation of age and sex of preadolescent children from bones and teeth. *Amer. J. phys. Anthrop.,* 1955, 13, 479-488.
5. NANDA, R. S. Agenesis of the third molar in man. *Amer. J. Orthodont.,* 1954, 40, 698-706.
6. PYLE, S. I., MANN, A. W., DREIZEN, S., KELLY, H. J., MACY, I. G., & SPIES, T. D. A substitute for skeletal age (Todd) for clinical use: The red graph method. *J. Pediat.,* 1948, 32, 125-136.

7. Saito, H. Röntgenologische Untersuchungen über die Entwicklung des dritten Molaren. *Kakubyo-Gakkai-Zasshi,* 1936, 10, 156-171, 366-377, 502-514.

8. Steggerda, M. Anthropometry and the eruption time of teeth. *J. Amer. dent. Ass.,* 1945, 32, 339-342.

9. Talmers, D. A. Time of eruption of second permanent molar and relationship to body size and areolar development. Preliminary report. *N. Y. State dent. J.,* 1952, 18, 314-315.

EAU OF EDNL. &

EDYL. & PSYL. RESEARCH
LIBRARY
21 JAN 1966